12.X.15 5/6 net. 66

MATHEMATICAL MONOGRAPHS

EDITED BY

MANSFIELD MERRIMAN AND ROBERT S. WOODWARD

No. 15

MORTALITY
LAWS AND STATISTICS

BY

ROBERT HENDERSON,
ACTUARY OF THE EQUITABLE LIFE ASSURANCE SOCIETY OF THE UNITED STATES

FIRST EDITION
FIRST THOUSAND

NEW YORK
JOHN WILEY & SONS, INC.
LONDON: CHAPMAN & HALL, LIMITED
1915

THE SCIENTIFIC PRESS
ROBERT DRUMMOND AND COMPANY
BROOKLYN, N. Y.

PREFACE

THE present work is designed to set forth in concise form the essential facts and theoretical relations with reference to the duration of human life. A description is first given of those mortality tables which have had the greatest influence on the development of the science of life contingencies or on its application in this country. A few chapters are then devoted to the mathematical relations between the various functions connected with human mortality, to the analysis of probabilities of death or survival, so as to lead to their simplest form of expression in terms of the mortality table, and to the general mathematical laws which have been proposed to express the facts of human mortality. The connection is then established between the mortality table and mortality statistics and some investigation made of the corrections which must be allowed for in interpreting such statistics.

The methods of constructing mortality tables from census and death returns and from insurance experience are then taken up. The methods adopted for the purpose of adjusting the rough data derived from experience are next described and their theoretical basis investigated. Some of these methods of construction and graduation are then illustrated by a new mortality table now first published. In the Appendix ten useful tables are given.

The scope of the treatise is confined to life contingencies excluding all monetary applications, so that the combination of the theory of compound interest with that of life contingencies is not touched upon. A warning may not, however, be amiss that the present value of a sum of money payable

at death cannot properly be calculated by assuming it to be payable at the end of a definite period equal to the expectation of life, nor can the present value of a life annuity be calculated by assuming it to be certainly payable for that period.

R. HENDERSON.

NEW YORK, March 1, 1915

CONTENTS

MORTALITY TABLES

DIAGRAMS

MORTALITY LAWS AND STATISTICS

CHAPTER I

MORTALITY TABLES

1. THE subject of human mortality is one which, from its nature, is of widespread interest to mankind. It has always been recognized that it is impossible to predict the duration of any individual life and that the only thing that could be taken as certain on the subject was that death would come sometime to each one. In other words it has been recognized that the date of death of any individual is subject to chance. The scientific study of the subject of chance is, however, a comparatively modern development of mathematics and consequently the science of life contingencies is also comparatively modern.

2. Like all other events, whether considered as chance events or as certainties, the death of any individual or his survival to any specified date is the necessary result of those forces which have been operating upon him. The cause of our ignorance regarding the result in the case of an individual is the limitation of our knowledge regarding the forces operating and their effects. We do know, however, that among the important ones affecting the result are climate, sanitary conditions, medical attendance and habits of life. These vary in their tendency and effectiveness as we pass from one locality to another or at different times in the same locality and may even differ in a recognizable way as between different individuals in the same locality and at the same time. The results of the observations made with respect to human mor-

tality under any given set of circumstances are frequently set forth concisely in the form of a *Mortality Table* showing the number surviving to each age out of a given number living at some selected initial age. A brief description is here given of some of the Mortality Tables which have had a relatively important part in the history of the science of life contingencies.

The Breslau Table

3. This table is of importance because it represents the first attempt to construct a mortality table from which to deduce the probabilities of survival and the values of life annuities. It was formed by the celebrated astronomer, Dr. E. Halley, from returns of the deaths in the City of Breslau, Silesia, during the five years 1687 to 1691 inclusive. Owing to the fact that the births during the five years only slightly exceeded the deaths (the numbers being 6193 and 5869 respectively) he assumed that the population might be considered a stationary one. He therefore appears to have graduated by inspection the average number dying per annum at the various ages and assumed that this gave the decrements of the table. Dr. Halley published his table in two columns, the first headed "Age Current," and the second "Persons," and it appears from the explanation given to have been in modern notation equivalent to a table of values of L_{x-1}, the population at age x next birthday.

4. On the basis of this table Dr. Halley solved various problems regarding survival and calculated annuity values, thus laying the foundations of the science of life contingencies and preparing the way for the transaction, on a scientific basis, of the important business of life insurance, although it was not until nearly seventy years had elapsed after the publication in 1693 of this table that the first company to operate on a scientific basis was established. Mr. E. J. Farren, writing in 1850, said regarding this table:

"With respect to its form, as has already been stated, no improvement has as yet been adopted, beyond inserting

the column of differences or deaths, and choosing higher numbers for exemplification. Of its two principles of construction, viz., as to the number of living being deducible from the number of deaths, by aid of the assumption of a stationary population; and as to the number of deaths at contiguous ages after childhood being allied in number; the former principle was generally prevalent in the construction of such tables, until the appearance of Mr. Milne's Carlisle Table in 1815, but is now as generally abandoned; the latter characteristic is still operative and considered as valid in all the best tables."

The mortality indicated by this table was considerably higher than that shown by more modern tables.

5. The table as given by Dr. Halley is as follows:

Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.
1	1000	8	680	15	628	22	586	29	539	36	481
2	855	9	670	16	622	23	579	30	531	37	472
3	798	10	661	17	616	24	573	31	523	38	463
4	760	11	653	18	610	25	567	32	515	39	454
5	732	12	646	19	604	26	560	33	507	40	445
6	710	13	640	20	598	27	553	34	499	41	436
7	692	14	634	21	592	28	546	35	490	42	427

Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.	Age Current.	Persons.
43	417	50	346	57	272	64	202	71	131	78	58
44	407	51	335	58	262	65	192	72	120	79	49
45	397	52	324	59	252	66	182	73	109	80	41
46	387	53	313	60	242	67	172	74	98	81	34
47	377	54	302	61	232	68	162	75	88	82	28
48	367	55	292	62	222	69	152	76	78	83	23
49	357	56	282	63	212	70	142	77	68	84	20

Age.	Persons
7	5 547
14	4 584
21	4 270
28	3 964
35	3 604
42	3 178
49	2 709
56	2 194
63	1 694
70	1 204
77	692
84	253
100	107
Sum	
Total	34 000

The column on the right hand is evidently a summary of the preceding figures in groups of seven years; with an additional item giving 107 as the population at ages 85 to 100 inclusive.

The Northampton Table

6. This table was first published in 1783 in the fourth edition of Dr. Price's work on Reversionary Payments. It was constructed by Dr. Price on the basis of a return of the deaths in the parish of All Saints, Northampton, England, during the forty-six years from 1735 to 1780 inclusive. These deaths by ages were as follows:

Ages.	Number of Deaths.
0–2	1529
2–5	362
5–10	201
10–20	189
20–30	373
30–40	329
40–50	365
50–60	384
60–70	378
70–80	358
80–90	199
90–100	22
Total.........	4689

7. Owing to the fact that the total number of baptisms during the same period was only 4220, or almost exactly ten per cent less than the number of deaths, Dr. Price, apparently believing the baptisms to correctly represent the births, assumed a stationary population supported by births and immigration at age 20. He supposed that the immigration was sufficient to supply thirteen per cent of the total deaths. The actual process followed appears to have been to transfer a sufficient number of deaths from age group 20 to 30 to age group 30 to 40 to equalize the numbers in the two groups. The number of deaths in the various groups was then proportionately increased so as to make the total number 10 000. Thirteen per cent of this number or 1300 in all were then deducted pro rata from the groups above age 20. All the groups both below and above age 20 were then increased pro rata so as to bring the deaths above age 20 up to the same figure

as before. This increased the total deaths to 11 650, which was taken as the value of l_0. A subtraction of the deaths in the successive groups gave the values of l_2, l_5, l_{10}, l_{20}, etc., the intermediate values being aftcrwards inserted.

8. From the above explanation it will be seen that in this case, as in the case of the Breslau table, only the death returns were available, without any enumeration of the population, and the resulting table indicated rates of mortality, especially at the younger ages, which, in the light of subsequent experience, appear unduly high. In fact, instead of being stationary, the population and the births had been regularly increasing and the population and deaths at the young ages were consequently proportionately higher than would have been the case in a stationary population. Thus when it was assumed that the number attaining any given age was equal to the number dying above that age the result was an understatement of the former number by the amount of the total increase during that period in the population above that age, subject to adjustment for immigration or emigration. This understatement of the denominator of the fraction determining the rate of mortality, of course, overstated that rate.

9. The Northampton Table was adopted by the Equitable Society as a basis for its calculations immediately after its construction. This fact, combined with the success of that Society, caused its adoption for many purposes for which it was not suitable. An outstanding illustration of this is the fact that the British Government based their rates for the sale of annuities upon it and consequently sustained a serious loss because the longevity of the annuitants proved much greater than was indicated by the table. Until within a few years the Northampton Table with five per cent interest was the basis prescribed by the court rules in New York State for the valuation of life interests and dower rights, but the Carlisle Table has now been adopted instead.

The Carlisle Table

10. This table was constructed in 1815 by Dr. Milne and was the first table to take both the deaths and the corresponding population into account. It was based on two censuses of the population of the parishes of St. Mary and St. Cuthbert, Carlisle, taken January 1, 1780, and December 31, 1787, or an interval of eight years, and on the deaths for the nine years 1779 to 1787.

The following schedule shows the data:

Age Last Birthday.	Population.		Deaths 1779 to 1787.
	Jan. 1780.	Dec. 1787.	
0			390
1			173
2	1029	1164	128
3			70
4			51
5– 9	908	1026	89
10– 14	715	808	34
15– 19	675	763	44
20– 29	1328	1501	96
30– 39	877	991	89
40– 49	858	970	118
50– 59	588	665	103
60– 69	438	494	173
70– 79	191	216	152
80– 89	58	66	98
90– 99	10	11	28
100–104	2	2	4
Total.......	7677	8677	1840

11. The figures for the deaths are derived from a record kept by Dr. J. Heysham, and the population as of January, 1780, is also derived from an enumeration by him, taking account of the ages. The population as of December, 1787, appears to have been merely enumerated in gross and then distributed by ages in the same proportions as had been found in 1780. It will be seen on examination that the proportions on the two dates are the same, which is a condition scarcely likely to be realized in two actual enumerations. It was assumed

that the average population during the period covered by the observations could be represented by the mean of the two censuses. The method followed in deducing the rates of mortality will be described in Chapter VI.

This table presented a more accurate statement of the probabilities of death at various ages than any preceding table and was widely used for insurance calculations. It has now been largely superseded for this purpose by tables more recently constructed from the experience of insured lives. It is, however, still in use for special purposes.

Owing to the small extent of the data on which it was based and the graphic method adopted in redistributing the population and deaths into individual ages, the rates of mortality were somewhat irregular, particularly at the older ages, and various regraduations of the table have been made with the idea of removing the irregularities.

The English Life Tables

12. At various times tables of mortality have been constructed on the basis of the census returns and registration of deaths in England. On account of the fact that an extensive series of monetary tables was based on it the most widely known of these tables is the English Life Table No. 3, which was constructed by Dr. Farr on the basis of the censuses of 1841 and 1851 and the deaths of the seventeen years 1838–54. Separate mortality tables were constructed for male and female lives starting with radices of 511 745 and 488 255 respectively or for the combined table, 1 000 000 at age zero. The method followed in constructing these tables will be described in Chapter VI.

13. At about the same time as these tables were constructed other tables, known as the Healthy Districts Life Tables, were also constructed from the census returns for 1851 of the sixty-four English districts having at that time an average death rate below 17 per thousand, and the deaths in the same districts during the five years 1849–53. These tables were pre-

sented in threefold form, the radix at age zero for the male table being 51 125 and for the female 48 875, the two added together constituting a mixed or combined table with a radix 100 000. The Healthy Districts Male Table with certain modifications was used by the Committee of the Actuarial Society of America in charge of the Specialized Mortality Investigation as a basis for the comparison of the mortality in the different classes.

14. The more recent of the series of English Life Tables are designated as Nos. 6, 7, and 8. The English Life Tables No. 6 were based on the census returns of 1891 and 1901 and the deaths of the ten years 1891 to 1900 inclusive and, while not originally prepared by Mr. Geo. King's method described in Chapter VII, have been readjusted by that method. The English Life Tables Nos. 7 and 8 have just been published and were prepared by Mr. Geo. King by his method. The former set are based on the census returns of 1901 and 1911 and the deaths of the ten years 1901 to 1910 inclusive, while the latter are based on the census of 1911 adjusted for increase to the middle of that year and on the death returns of the three years 1910 to 1912 inclusive. The special feature of these tables is that not only do they indicate an improvement in mortality as compared with the earlier tables of the series, but, when compared with one another they indicate that the improvement was still progressing. The No. 7 Tables show a lower mortality throughout than the No. 6 and the No. 8 Tables a lower mortality at practically all ages than the No. 7.

The Actuaries', or Combined Experience, Table

15. This table, also known as the Seventeen Offices' Experience Table, was prepared in 1841 by combining the experience, by lives, of the Equitable and Amicable Societies with the experience, by policies, of fifteen other companies as contributed in 1838 to a committee of actuaries. It was thus the first example of a mortality table formed by combining the experiences of different insurance companies into one

general average. It appears to have covered in all 83 905 policies or lives of which 13 781 were terminated by death, 25 247 were terminated otherwise, and 44 877 were in existence and under observation when the observations closed. The total of the numbers exposed to risk, for one year at each age, was 712 163 indicating an average duration of 8.5 years.

16. Probably owing to the mixed nature of the data, which as above stated, was partly by lives and partly by policies, and to the fact that the average duration of the experience contributed by the companies other than the Equitable and the Amicable was only 5.5 years, this table was never widely used in Great Britain for insurance purposes. It was, however, prescribed by the State of Massachusetts as the basis for the valuation of the reserve liabilities of life insurance companies. The example of Massachusetts was later followed by New York and other states with the result that for many years the Actuaries' Table with four per cent interest was the accepted valuation standard in the United States, although the premiums actually charged by the companies were as a rule based on a different table.

The Healthy Male (H^M) Table

17. This is the most important of the group of tables published in 1869 and representing the results of the Institute of Actuaries' Mortality Experience, 1863. They were based upon data contributed by twenty British life insurance companies regarding their experience up to 1863 on insured lives. The H^M table was based on the experience of male lives insured at regular premium rates, and duplicate policies on the same life, whether in the same or in different companies, were carefully eliminated. This table represented a much broader experience than that upon which the Actuaries' Table had been based, confirmed in a general way the results of that experience and obtained immediate acceptance as a fair representation of the average mortality of insured lives. The official H^M Table was graduated by Woolhouse's formula

but it was subsequently regraduated by King and Hardy according to Makeham's formula, with a modification at the younger ages, and extended down to age zero by means of rates of mortality taken from the Healthy Districts Male Table. This graduation of the table is published in Part II of the Text Book of the Institute of Actuaries.

18. In the construction of this table all lives of the same attained age were included together without regard to the period elapsed since medical examination. But an analysis of the experience indicated that the rate of mortality among lives recently insured was much less than among lives of the same attained age who had been insured for a longer period. Accordingly a second table, known as the $H^{M(5)}$ Table, was formed by omitting the experience during the calendar year of issue and the next four calendar years. This table was taken as representing the ultimate rate of mortality after the effects of selection had worn off. The rates of mortality at the young ages are considerably higher in the $H^{M(5)}$ table than in the H^M, but the two rates gradually approach one another and coincide at the extreme old ages where there are no recently selected lives.

19. These two tables used together were adopted by many British companies for the valuation of their liabilities, and the H^M Table was for many years prescribed for that purpose by the laws of Canada. It will be noticed that the rates of mortality according to the H^M Table are lower than those for the same ages in the Actuaries' Table except for ages 46 to 50 and ages 73 to 85 inclusive and ages 95 and over. The difference is not, however, important except at the young ages, where it is considerable.

The British Offices' Life Tables, 1893

20. These tables represent the experience on insured lives of sixty British life insurance companies during the thirty years from the policy anniversaries in 1863 to those in 1893. The data were compiled under the joint supervision of the

Institute of Actuaries and the Faculty of Actuaries in Scotland and was classified into male and female lives and according to the plan of insurance issued. The O^M Table represents the experience of male lives insured on the Ordinary Life plan with participation in profits. The total number of lives under observation was 551 838, of whom 149 566 were insured prior to 1863. Of these 140 889 died, 148 392 withdrew and 262 557 remained insured in 1893, the total number of years of risk being 7 056 863. The $O^{M(5)}$ Table represents the same experience, omitting the first five policy years and covers 5 324 862 years of risk and 129 001 deaths. These tables were graduated by Mr. G. F. Hardy. The $O^{M(5)}$ table was first graduated by the application of Makeham's formula, the differences in the values of log p_x by the two tables being then graduated by the use of a double-frequency curve. A select or analyzed table was also prepared from the same data and is known as the $O^{[M]}$ Table. In this table separate rates of mortality are indicated for each age at entry for the first ten policy years, merging into an ultimate table at the end of that time. This select table was also graduated by Makeham's formula, different constants being used for the different policy years. The rates of mortality by the O^M table are lower throughout than those in the H^M table as graduated by Makeham's formula and also, with unimportant exceptions, than those in the official H^M table. The $O^{M(5)}$ Table is the basis at present prescribed for the valuation of policies in Canada.

The American Experience Table

21. This table was constructed by Mr. Sheppard Homans and was first published in its present form in 1868. No complete record has ever been made public of the method adopted in its construction, but it has always been understood that the mortality experience of the Mutual Life Insurance Company of New York was used as a basis. As that experience covered only a few years and therefore did not include any exposures or deaths at extreme old ages it must have been supplemented

from other sources. The table is a very smoothly graduated one and evidence has been discovered which seems to indicate that the author first constructed a table of values of the reciprocal of the rate of mortality showing the number of lives out of which one death would be expected at each age. From these values the usual columns of the mortality table were then formed.

22. The first publication of the table was in the schedule of an act prescribing it as a basis of valuation in the State of New York and although it was temporarily abandoned in that state for the sake of uniformity it is now the legal standard in practically every state of the Union. The table as originally published was found to conform very nearly to Makeham's law, and was subsequently regraduated in accordance with that law for use in connection with joint life calculations.

The American Experience Table has been widely used in America as a basis for insurance premiums even when another table was prescribed as a legal basis of valuation, as it presented a conservative view of the mortality after the effect of selection had worn off. The rates of mortality shown were higher than those in the Actuaries' table for ages 30 and under and for ages 78 and over, but lower between 30 and 78. Compared with the H^M Table, which was published about the same time, it gave higher rates of mortality for ages under 36 and over 80 and slightly lower for the intervening years. Compared with the $O^{M(5)}$ it shows higher rates of mortality for ages 40 and under and for ages over 70 and lower values for the intermediate ages. It will thus be seen that in general the American Experience Table seems to give relatively low rates of mortality for the central ages and high rates for the young and old ages.

The National Fraternal Congress Table

23. This table was constructed by the Committee on Rates of the National Fraternal Congress, an association of Fraternal Societies in the United States of America, and was presented

in its original form at the annual meeting of that association in 1898. It was based on the experience up to that time of the societies connected with the Congress. It was subsequently regraduated by Mr. Abb Landis and reported in its amended form the next year. Compared with the American Experience Table the rates of mortality are lower throughout, although the difference is proportionately smaller at the older ages than at the younger. Compared with the O^M Table the rates of mortality are higher at ages 20 to 27 inclusive and for ages 81 and over and lower at the intervening ages.

24. This table of mortality with interest at four per cent is prescribed as a basis for minimum rates of contribution in fraternal orders by the laws of several States. It is worthy of note that a subsequent investigation was made of the experience during the year 1904 of 43 societies. This experience covered 2 880 166.5 years of exposure and 19 414 deaths, and the rates of mortality in the resulting table were lower than those in the National Fraternal Congress Table for ages up to 52 inclusive and for ages 79 and over, but higher for the intervening ages.

The M. A. Table of the Medico-Actuarial Mortality Investigation

25. This table was constructed in 1912 by the joint committee of the Medical Directors' Association and the Actuarial Society of America in charge of the Medico-Actuarial Mortality Investigation into the relative mortality of special classes of risks. It was intended for use as a standard with which to compare the mortality of the special classes. It was therefore based on the experience of the same companies as contributed to the special class experience on policies issued during the same period and observed up to the same date. The data used were based on the experience of the companies on policies issued during the month of January in odd years and July in even years from 1885 to 1908 inclusive, observed to the anniversaries in 1909. The total number of policies was 500 375,

the total years of exposure 2 814 276 and the number of policies terminated by death 20 222. The table is shown in the form of analyzed rates of mortality for the first four policy years with an ultimate table for the fifth and subsequent years.

A special feature of this table is that the difference between the rates of mortality in the early policy years and those shown for the same attained ages in the ultimate table is relatively small. This has been explained on the theory that an improvement in general mortality conditions was going on during the time of the observations and that, owing to the fact that the observations in the early policy years were on an average made at an earlier date than those for the longer durations, this partly concealed the true effect of selection. This theory was confirmed by investigating separately the experience on policies issued in the years 1885 to 1892 inclusive, those issued in 1893 to 1900 inclusive, and those issued in 1901 to 1908 inclusive. A progressive improvement was shown in passing from one group to the next. In the ultimate part of the M. A. Table the rates of mortality are throughout lower than those for the same age in the American Experience Table, but practically equal at age 69. Compared with the National Fraternal Congress Table, they are lower at ages under 55 and over 80 but higher between those ages. For ages under 70 the rates of mortality are lower than in the ultimate part of the British Offices' $O^{[M]}$ Table, but after that age they agree exactly with that table.

26. This table was constructed in a special way for the special purpose above indicated and is not recommended by its authors for any other purpose. The experiences, however, of some individual companies which have since been investigated appear to confirm substantially the ultimate part of the table as a fair representation of ultimate mortality of insured lives in American and Canadian Companies transacting a normal business.

McClintock's Annuitants' Mortality Tables

27. These tables were constructed in 1899 by Dr. McClintock on the basis of experience of fifteen American companies, collected and analyzed by Mr. Weeks. The data comprised the entire experience of the companies on annuities up to the anniversaries of the contracts in 1892. Separate tables were constructed for male and female lives, the number of lives taken into consideration being 4365 males and 4821 females. Although this was an experience of American companies only about one-fourth of the number of annuitants were actually American lives, the remaining three-fourths representing annuities granted abroad by the companies. The experience was taken out strictly by lives, all duplicates being carefully eliminated, and in the case of deferred annuities only the experience after the annuity became payable was considered, owing to the uncertainty with regard to the date of death during the deferred period.

28. Each table was graduated by Makeham's formula, (Art. 50), the same value of c being used for the two tables and in consequence of this fact the principle of uniform seniority may be used, although in a modified form, even where the lives are not all of the same sex. The formula adopted was $\operatorname{colog} p_x = \log b + c^x \log h$, where $\log c = .04$ and for the male table $\log b = .0032$ and $\log \log h = \bar{5}.55$; for the female table $\log b = .0015$ and $\log \log h = \bar{5}.43$. The rate of mortality is higher throughout the male table than for the same age in the female table, the difference being proportionately greatest at the young ages. The rate of mortality in the male is higher than in the American Experience Table up to age 62 and lower above that age. In connection with these tables it should be remembered that at the young ages they are purely theoretical, there being only two actual deaths at ages under 40 in the male experience and three in the female. These tables are now prescribed by the law of New York State as the basis for the valuation of annuity contracts issued by life insurance companies.

The British Offices' Life Annuity Tables, 1893

29. These tables are derived from the experience of British Offices in respect of life annuitants, male and female, during the period 1863 to 1893, including the British Annuity experience of three American companies. Both select and aggregate unadjusted tables were constructed, duplicates being separately eliminated for each. After the final elimination of duplicates for the aggregate tables the total number of male lives involved was 6728, the number of years of risk 53 599 and the number of deaths 3503. For the female table the number of lives was 18 951, the number of years of risk 173 519 and the number of deaths 9107.

30. The graduated tables constructed from these data were shown in the select or analyzed form with separate rates of mortality for each of the first five contract years, merging into an ultimate table at the end of the fifth year. The male table was graduated by Makeham's formula (Art. 50), modified for duration, the value of $\log_{10}c$ being .038. The female table could not be graduated as a single series by that law. A second series was therefore introduced and it was assumed that $l_{[x]+t}=l^{(1)}_{[x]+t}+l^{(2)}_{[x]+t}$, where $l^{(1)}_{[x]+t}$ and $l^{(2)}_{[x]+t}$ each conformed to Makeham's formula modified for duration. The rates of mortality in the ultimate part of the male table are lower than in McClintock's table for ages under 50 and over 82 and higher for the intervening ages. In the ultimate female table the rates of mortality are higher throughout than in McClintock's table. The value at $3\frac{1}{2}$ per cent interest of an annuity at date of issue is somewhat higher by the British Offices' Male Table throughout than by McClintock's Table. By the female table the value at date of issue is lower than by McClintock's table for ages under 62 and from age 69 to age 75 inclusive and higher for ages 62 to 68 inclusive and for ages over 75.

CHAPTER II

THE MORTALITY TABLE AND PROBABILITIES INVOLVING ONE LIFE

31. The mortality table has been defined as "the instrument by means of which are measured the probabilities of life and the probabilities of death." It may be considered as primarily a table showing how many on an average survive to each attained age out of a given number living at some selected initial age. The symbol l_x is used to denote the number surviving to age x and if a be the initial age selected it is evident that l_a represents the given number observed, since they are all living at that time. The mortality table, therefore, asserts that on the average out of l_a persons living at age a, l_x will survive to age x, where x is any higher age. But if on the average in each N out of a series of cases in which an event A is in question A happens on pN occasions, the probability of the event A is said to be p. The probability, therefore, that a life aged a will survive to age x is l_x/l_a.

32. This is a property, however, which is not confined to the initial age a. Consider any third age y greater than x. The probability, then, of a life aged a surviving to age y will be l_y/l_a. But this event may be considered as a compound event, being composed of a life aged a surviving to age x and a life aged x surviving to age y. The probability of the first is l_x/l_a; therefore, by division, the probability of the second is l_y/l_x. This may also be demonstrated from the consideration that the l_y survivors at age y are the survivors out of l_x living at age x, because they are all included among the l_x, and there are none included in l_x who, if surviving at age y, would not be included in l_y. Therefore, again the probability of a life aged x surviving to age y is l_y/l_x. A single table,

therefore, of the values of l_x gives by a single division the probability of survival for any age and period included in its range.

33. The probability of a life aged x surviving n years is designated by ${}_np_x$, and since the attained age at the end of the period is $x+n$ we have

$$ {}_np_x = l_{x+n}/l_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1) $$

The probability of surviving one year is denoted by p_x, so that

$$ p_x = l_{x+1}/l_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2) $$

From these equations it is evident that we have the general relation

$$ {}_np_x = p_x \cdot p_{x+1} \cdot p_{x+2} \cdot \cdot \cdot p_{x+n-1}. \quad . \quad . \quad . \quad (3) $$

We thus see that from a complete table of the values of p_x the probabilities over longer periods can be calculated. Eq. (2) can in fact be stated in the form $l_{x+1} = l_x p_x$ by the application of which the successive values of l_x can be calculated, starting from any given value.

34. Hitherto we have dealt with the probability of survival over a specified period. The complementary probability is that of death within the period. The probability of a life aged x dying within one year is denoted by q_x. Since the life must either survive one year or die within the year we have

$$ p_x + q_x = 1 \quad \text{or} \quad q_x = 1 - p_x, $$

whence we obtain the following

$$ q_x = (l_x - l_{x+1})/l_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4) $$

It is usual to designate the function $(l_x - l_{x+1})$ by d_x, so that d_x denotes the number dying between the ages x and $x+1$, or at age x last birthday, out of l_x living at age x; or out of l_a living at age a.

35. Suppose that $w-1$ is the highest age at which any survivors are recorded in the mortality table, so that $l_{w-1} = d_{w-1}$ and $l_w = 0$, then

$$d_x+d_{x+1}+\ldots+d_{w-2}+d_{w-1}=(l_x-l_{x+1})+(l_{x+1}-l_{x+2})+\ldots +(l_{w-2}-l_{w-1})+l_{w-1}=l_x, \quad .$$

or

$$l_x=\Sigma_x^{w-1}d_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

In fact, if a frequency curve is supposed to be drawn such that the height of the ordinate corresponding to any age x is proportionate to the number dying at that age, then d_x will be the area included between the ordinates for ages x and $x+1$, and l_x will be the entire area between the ordinate for age x and the limit of the curve.

36. The probability of a life aged x dying within n years is denoted by $|_n q_x$ and we have, since $|_n q_x+{}_np_x=1$,

$$|_n q_x=1-{}_np_x=(l_x-l_{x+n})/l_x, \quad . \quad . \quad . \quad . \quad . \quad (6)$$

37. The probability that a life aged x will die in the nth year from the present time is evidently compounded of the probability that it will survive $n-1$ years and that having done so it will then die within one year, and is denoted by ${}_{n-1}|q_x$. Hence we have ${}_{n-1}|q_x={}_{n-1}p_x\cdot q_{x+n-1}$.

From this equation or from the consideration that the number, out of l_x living at age x, who die in the nth year thereafter is d_{x+n-1} we have

$${}_{n-1}|q_x=\frac{l_{x+n-1}}{l_x}\cdot\frac{d_{x+n-1}}{l_{x+n-1}}=\frac{d_{x+n-1}}{l_x}. \quad . \quad . \quad . \quad . \quad . \quad (7)$$

38. We have heretofore considered only the values of l_x for integral values of x, but it is evident that deaths occur at all times throughout each year of age so that l_x may be considered as a continuously varying function. Let us investigate its rate of decrease at any particular age. This rate is the limit when n vanishes of the function $(l_x-l_{x+n})/n$ which represents the average number dying per annum over a period of n years. This limit may be expressed in the language of the infinitesimal calculus as $-\frac{dl_x}{dx}$. The ratio of this instantaneous rate of decrease of l_x to the corresponding value of l_x

is called the force of mortality and is denoted by μ_x, so that we have

$$\mu_x = -\frac{dl_x}{dx}/l_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Unless, however, it is possible to express l_x as an algebraic function of x we cannot determine exactly the value of $\frac{dl_x}{dx}$, and consequently of μ_x. Certain approximate expressions can, however, be determined on the assumption that differential coefficients of a high order may be neglected. We have

$$l_{x+h} = l_x + h\frac{dl_x}{dx} + \frac{h^2}{2}\frac{d^2l_x}{dx^2} + \frac{h^3}{6}\frac{d^3l_x}{dx^3} + \frac{h^4}{24}\frac{d^4l_x}{dx^4} + \text{etc.},$$

and

$$l_{x-h} = l_x - h\frac{dl_x}{dx} + \frac{h^2}{2}\frac{d^2l_x}{dx^2} - \frac{h^3}{6}\frac{d^3l_x}{dx^3} + \frac{h^4}{24}\frac{d^4l_x}{dx^4} + \text{etc.},$$

by Taylor's theorem, whence,

$$l_{x+h} - l_{x-h} = 2h\frac{dl_x}{dx} + \frac{h^3}{3}\frac{d^3l_x}{dx^3} + \text{etc.}$$

If we assume that $\frac{d^3l_x}{dx^3}$ and higher orders may be neglected, and put $h=1$, we have as a first approximation

$$\frac{dl_x}{dx} = \tfrac{1}{2}(l_{x+1} - l_{x-1}),$$

whence

$$\mu_x = -\frac{dl_x}{dx}/l_x = (l_{x-1} - l_{x+1})/2l_x. \quad . \quad . \quad . \quad . \quad . \quad (9)$$

39. Next let us assume that $\frac{d^5l_x}{dx^5}$ and higher orders may be neglected, and substitute for h successively 1 and 2. Then we have

$$l_{x+1} - l_{x-1} = 2\frac{dl_x}{dx} + \frac{1}{3}\frac{d^3l_x}{dx^3},$$

$$l_{x+2} - l_{x-2} = 4\frac{dl_x}{dx} + \frac{8}{3}\frac{d^3l_x}{dx^3},$$

whence

$$8(l_{x+1}-l_{x-1})-(l_{x+2}-l_{x-2})=12\frac{dl_x}{dx},$$

and

$$\mu_x=\{8(l_{x-1}-l_{x+1})-(l_{x-2}-l_{x+2})\}/12l_x. \quad . \quad . \quad (10)$$

40. It will be noted that by the principles of the infinitesimal calculus we have

$$\frac{d\log_e l_x}{dx}=\frac{dl_x}{dx}/l_x,$$

whence we have,

$$\mu_x=-\frac{d\log_e l_x}{dx}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

41. We have already pointed out that the distribution of the lives according to duration may be represented by a frequency curve, and in this case as in the case of frequency curves in general, we may seek for some typical value to represent the curve as a whole. There are three quantities sometimes used for this purpose. The first is the mode, or that value of the variable for which the probability is the greatest. In the mortality table this corresponds to the age at death for which $l_x\mu_x$ is the greatest. The most probable duration of life is therefore the difference between the present age and this definite age provided the present age is less than the age of maximum deaths.

42. The second quantity corresponds to the median in the theory of frequency curves and is the duration which the life has an even chance of surviving. It is known as the *vie probable* sometimes translated into probable lifetime. Its value for any age x is determined by solving for n the equation

$${}_np_x=\tfrac{1}{2} \quad \text{or} \quad l_{x+n}=\tfrac{1}{2}l_x.$$

43. The function most commonly used, however, for the purpose of summarizing the probabilities of survival of a given life is the expectation of life which corresponds to the mean value in the theory of frequency curves. The curtate expectation of life is the average or expected number of com-

plete years survived by lives of a given age. Its value may be calculated as follows:

Out of l_x lives at age x, d_x die in the first year without completing a year of life after age x, d_{x+1} die in the second year after completing one year, d_{x+2} after completing two years and so on. Therefore if we designate the curtate duration of life at age x by e_x we have,

$$l_x e_x = d_{x+1} + 2d_{x+2} + 3d_{x+3} + \text{etc.} \quad . \quad . \quad . \quad (12)$$

Substituting now for d_{x+1}, etc., their values in terms of l_x, we have

$$l_x e_x = (l_{x+1} - l_{x+2}) + 2(l_{x+2} - l_{x+3}) + 3(l_{x+3} - l_{x+4}) + \text{etc.},$$

$$= l_{x+1} + l_{x+2} + l_{x+3} + \text{etc.},$$

or

$$e_x = \frac{l_{x+1} + l_{x+2} + l_{x+3} + \text{etc.}}{l_x} = \frac{\Sigma_{n=1}^{n=\infty} l_{x+n}}{l_x}, \quad . \quad . \quad . \quad (13)$$

$$= \Sigma_{n=1}^{n=\infty} \frac{l_{x+n}}{l_x} = \Sigma_{n=1}^{n=\infty} {}_n p_x \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

This same expression for the curtate expectation may also be obtained by considering separately the number completing each year. The number completing the first year is l_{x+1}, those completing the second year l_{x+2}, the third year l_{x+3}, and so on. Therefore the total number of years completed by the l_x lives is $l_{x+1} + l_{x+2} + l_{x+3} + \text{etc.}$, and the average is obtained as above.

44. The calculation of the curtate expectation takes account only of years of life entirely completed and omits the fraction of a year survived in the year in which death occurs. The complete expectation of life, denoted by $\mathring{e}_x$, includes this fraction of a year. In arriving at a first approximation to the value of $\mathring{e}_x$, it is usual to assume that this fraction, which may have any value from zero to one year, averages half a year, so that we have

$$\mathring{e}_x = \tfrac{1}{2} + e_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

Or substituting for e_x its value in terms of l_x, we have

$$\mathring{e}_x=\frac{\frac{1}{2}l_x+l_{x+1}+l_{x+2}+l_{x+3}+\text{etc.}}{l_x},$$

$$=\frac{\frac{1}{2}(l_x+l_{x+1})+\frac{1}{2}(l_{x+1}+l_{x+2})+\frac{1}{2}(l_{x+2}+l_{x+3})+\text{etc.}}{l_x},$$

$$=\frac{L_x+L_{x+1}+L_{x+2}+\text{etc.}}{l_x}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

where generally,

$$L_x=\tfrac{1}{2}(l_x+l_{x+1}).$$

45. The exact expression for the complete expectation of life is evidently

$$l_x\mathring{e}_x=\int_0^\infty tl_{x+t}\mu_{x+t}dt=-\int_0^\infty t\frac{dl_{x+t}}{dt}\cdot dt. \quad . \quad . \quad (17)$$

Integrating by parts in the regular way this becomes

$$l_x\mathring{e}_x=-\left[tl_{x+t}\right]_0^\infty+\int_0^\infty l_{x+t}dt=\int_0^\infty l_{x+t}dt, \quad . \quad . \quad . \quad (18)$$

since tl_{x+t} vanishes at both limits. This integral cannot be evaluated exactly unless an algebraic expression in integrable form can be substituted for l_{x+t}. An approximate expression can, however, be obtained by dividing the integration into yearly intervals. We have

$$\int_t^{t+1} l_{x+t}dt=\int_0^1 l_{x+t+h}\cdot dh$$

$$=\int_0^1\left(l_{x+t}+h\frac{dl_{x+t}}{dt}+\frac{h^2}{2}\frac{d^2l_{x+t}}{dt^2}+\frac{h^3}{6}\frac{d^3l_{x+t}}{dt^3}+\frac{h^4}{24}\frac{d^4l_{x+t}}{dt^4}+\text{etc.}\right)dh.$$

$$=l_{x+t}+\frac{1}{2}\cdot\frac{dl_{x+t}}{dt}+\frac{1}{6}\frac{d^2l_{x+t}}{dt^2}+\frac{1}{24}\frac{d^3l_{x+t}}{dt^3}+\frac{1}{120}\frac{d^4l_{x+t}}{dt^4}+\text{etc.}$$

from Taylor's theorem; also,

$$\tfrac{1}{2}(l_{x+t}+l_{x+t+1})=l_{x+t}+\frac{1}{2}\frac{dl_{x+t}}{dt}+\frac{1}{4}\frac{d^2l_{x+t}}{dt^2}+\frac{1}{12}\frac{d^3l_{x+t}}{dt^3}+\frac{1}{48}\frac{d^4l_{x+t}}{dt^4}+\text{etc.}$$

$$\left(\frac{dl_{x+t+1}}{dt}-\frac{dl_{x+t}}{dt}\right)=\frac{d^2l_{x+t}}{dt^2}+\frac{1}{2}\frac{d^3l_{x+t}}{dt^3}+\frac{1d^4l_{x+t}}{6\,dt^4}+\text{etc.}$$

$$\left(\frac{d^3l_{x+t+1}}{dt^3}-\frac{d^3l_{x+t}}{dt^3}\right)=\frac{d^4l_{x+t}}{dt^4}+\text{etc.}$$

Hence we have

$$\int_t^{t+1} l_{x+t}dt=\tfrac{1}{2}(l_{x+t}+l_{x+t+1})-\frac{1}{12}\left(\frac{dl_{x+t+1}}{dt}-\frac{dl_{x+t}}{dt}\right)$$
$$+\frac{1}{720}\left(\frac{d^3l_{x+t+1}}{dt^3}-\frac{d^3l_{x+t}}{dt^3}\right)-\text{etc.}$$

Putting then t successively equal to 0, 1, 2, 3, etc., we get

$$\int_0^1 l_{x+t}dt=\tfrac{1}{2}(l_x+l_{x+1})-\frac{1}{12}\left(\frac{dl_{x+1}}{dx}-\frac{dl_x}{dx}\right)$$
$$+\frac{1}{720}\left(\frac{d^3l_{x+1}}{dx^3}-\frac{d^3l_x}{dx^3}\right)-\text{etc.}$$

$$\int_1^2 l_{x+t}dt=\tfrac{1}{2}(l_{x+1}+l_{x+2})-\frac{1}{12}\left(\frac{dl_{x+2}}{dx}-\frac{dl_{x+1}}{dx}\right)$$
$$+\frac{1}{720}\left(\frac{d^3l_{x+2}}{dx^3}-\frac{d^3l_{x+1}}{dx^3}\right)-\text{etc.}$$

etc. . . . .

Whence, summing and remembering that at the upper limit l_{x+t} and all its differential coefficients may be assumed to vanish, we have

$$\int_0^\infty l_{x+t}dt=\tfrac{1}{2}l_x+\Sigma_1^\infty l_{x+t}+\frac{1}{12}\frac{dl_x}{dx}-\frac{1}{720}\frac{d^3l_x}{dx^3}+\text{etc.},$$
$$=\tfrac{1}{2}l_x+\Sigma_1^\infty l_{x+t}-\frac{1}{12}l_x\mu_x-\frac{1}{720}\frac{d^3l_x}{dx^3}+\text{etc.}$$

Hence, if we neglect the term $\frac{1}{720}\frac{d^3l_x}{dx^3}$ and all higher differential coefficients, we have

$$l_x\mathring{e}_x=\tfrac{1}{2}l_x+\Sigma_1^\infty l_{x+t}-\tfrac{1}{12}l_x\mu_x,$$
$$=\tfrac{1}{2}l_x+l_xe_x-\tfrac{1}{12}l_x\mu_x,$$

or

$$\mathring{e}_x=\tfrac{1}{2}+e_x-\tfrac{1}{12}\mu_x. \qquad (19)$$

This shows that the correction to the first approximation is approximately $-\frac{1}{12}\mu_x$. As the value of this correction is very small except at extreme old age it is usually neglected and the first approximation used for $\mathring{e}_x$.

46. From a table of the expectation of life it is possible to derive directly the corresponding rates of mortality, for we have

$$l_x e_x = l_{x+1} + l_{x+2} + l_{x+3} + \text{etc.},$$

$$l_{x+1}(1 + e_{x+1}) = l_{x+1} + l_{x+2} + l_{x+3} + \text{etc.},$$

so that we have

$$l_x e_x = l_{x+1}(1 + e_{x+1}),$$

or

$$e_x = \frac{l_{x+1}}{l_x}(1 + e_{x+1}) = p_x(1 + e_{x+1}), \quad . \quad . \quad . \quad . \quad (20)$$

whence,

$$p_x = \frac{e_x}{1 + e_{x+1}},$$

and

$$q_x = \frac{1 - (e_x - e_{x+1})}{1 + e_{x+1}}. \quad . \quad . \quad . \quad . \quad . \quad (21)$$

Eq. (21) gives the rate of mortality in terms of expectations and Eq. (20) gives a rule for determining e_x from p_x and e_{x+1}. By the successive application of this formula, beginning at the oldest age, the expectation of life for all ages may be computed from the rates of mortality without constructing the l_x column.

47. The value of μ_x may also be expressed in terms of complete expectations of life, for we have $l_x \mathring{e}_x = \int_x^\infty l_x dx$, from Eq. (18). Differentiating, then, with respect to x we have, after changing sign,

$$\mu_x l_x \mathring{e}_x - l_x \frac{d\mathring{e}_x}{dx} = l_x,$$

or

$$\mu_x \mathring{e}_x = 1 + \frac{d\mathring{e}_x}{dx} = 1 - \tfrac{1}{2}(\mathring{e}_{x-1} - \mathring{e}_{x+1}), \text{ approximately}$$

whence

$$\mu_x = \frac{1 - \frac{1}{2}(\mathring{e}_{x-1} - \mathring{e}_{x+1})}{\mathring{e}_x}.$$

CHAPTER III

FORMULAS FOR THE LAW OF MORTALITY

48. Before the various labor-saving devices now in use in connection with the calculation of monetary values from the mortality table had been invented, the desirability of reducing, if possible, the mortality table to a mathematical law in order to facilitate such calculations was especially evident. The first attempt of this kind was made by DeMoivre in his "Treatise of Annuities on Lives," for the purpose of passing from the expectation of life to the value of a life annuity. His assumption was the very simple one that the value of d_x was the same for all ages, or in other words, that l_x decreased uniformly up to the limiting age. The equation for l_x in terms of x can therefore be written

$$l_x = a(w-x), \quad \ldots \ldots \quad (1)$$

where x is less than w.
Whence

$$\frac{dl_x}{dx} = -a = -d_x,$$

and therefore,

$$q_x = \mu_x = \frac{a}{a(w-x)} = \frac{1}{w-x}, \quad \ldots \ldots \quad (2)$$

also

$$l_x \mathring{e}_x = \int_x^w l_x dx = \int_x^w a(w-x)dx = -\frac{a}{2}\left[(w-x)^2\right]_x^w = \frac{a}{2}(w-x)^2$$

whence

$$\mathring{e}_x = \tfrac{1}{2}(w-x). \quad \ldots \ldots \quad (3)$$

This equation may also be stated in the form $w = x + 2\mathring{e}_x$, which shows that, if the formula applied, the function $x + 2\mathring{e}_x$ would be a constant. The calculation of this function for

two or three ages at intervals, or the examination of the d_x columns of any mortality table based on actual experience will show that DeMoivre's hypothesis is only a rough approximation to the truth. While it accomplished its purpose of enabling approximate life annuity values to be calculated from the expectation it cannot be accepted as a statement of the true law of mortality.

49. In the next attempt the problem was approached directly by an investigation of the causes of death. It was made by Benjamin Gompertz, who assumed that the force of mortality increases in geometrical progression with the age. This may be written as follows:

$$\mu_x = Bc^x, \quad \ldots \ldots \ldots \quad (4)$$

where B and c are constants for a given mortality table, but may have different values in different tables. From this equation we have

$$\frac{d \log_e l_x}{dx} = -Bc^x.$$

Whence, integrating with respect to x, we have

$$\log_e l_x = \log_e k - \frac{B}{\log_e c} c^x,$$

or

$$l_x = kg^{c^x}, \quad \ldots \ldots \ldots \quad (5)$$

where $\log_e g = -\frac{B}{\log_e c}$, or $B = -\log_e g \cdot \log_e c$. We may thus express μ_x in terms of the constants of the mortality table by substituting this value for B, so that we have

$$\mu_x = -(\log_e g \cdot \log_e c)c^x. \quad \ldots \ldots \quad (6)$$

50. Gompertz's formula constituted a genuine approximation to the law of mortality, but it was found that it did not apply to the period of childhood, and that even at adult ages it would not cover the complete range without a change of constant at an age in the neighborhood of 50 or 60. To remedy this Mr. Makeham proposed to modify Gompertz's formula

in a way actually suggested by the reasoning of Gompertz himself, who had stated that "It is possible that death may be the consequence of two generally coexisting causes; the one chance, without previous disposition to death or deterioration; the other a deterioration, or increased inability to withstand destruction." The modification consisted in adding a constant to the expression for the force of mortality, which became

$$\mu_x = A + Bc^x. \quad \ldots \ldots \ldots \quad (7)$$

$$= -\log_e s - (\log_e g \cdot \log_e c)c^x. \quad \ldots \quad (8)$$

by putting $A = -\log_e s$

Substituting then $-\dfrac{d \log_e l_x}{dx}$ for μ_x and integrating, we get

$$\log_e l_x = \log_e k + x \log_e s + c^x \log_e g,$$

where $\log_e k$ is the constant of integration, or

$$l_x = ks^x g^{c^x}. \quad \ldots \ldots \quad (9)$$

This formula has been applied to various mortality tables with considerable success, reproducing them very closely from about age 20 to the end, but not covering the period of infancy. Certain other tables, however, cannot be reproduced by this formula.

51. The simplest way of determining the constants in Makeham's formula is from four equidistant values of $\log l_x$. We have.

$$\log l_x = \log k + x \log s + c^x \log g,$$

$$\log l_{x+t} = \log k + (x+t) \log s + c^{x+t} \log g,$$

$$\log l_{x+2t} = \log k + (x+2t) \log s + c^{x+2t} \log g,$$

$$\log l_{x+3t} = \log k + (x+3t) \log s + c^{x+3t} \log g.$$

Taking now the differences, we have

$$\log l_{x+t} - \log l_x = t \log s + c^x(c^t - 1) \log g,$$

$$\log l_{x+2t} - \log l_{x+t} = t \log s + c^{x+t}(c^t - 1) \log g,$$

$$\log l_{x+3t} - \log l_{x+2t} = t \log s + c^{x+2t}(c^t - 1) \log g.$$

Taking differences again, we have

$$\log l_{x+2t} - 2 \log l_{x+t} + \log l_x = c^x(c^t - 1)^2 \log g,$$

$$\log l_{x+3t} - 2 \log l_{x+2t} + \log l_{x+t} = c^{x+t}(c^t - 1)^2 \log g.$$

Dividing the second by the first, we have

$$c^t = \frac{\log l_{x+3t} - 2 \log l_{x+2t} + \log l_{x+t}}{\log l_{x+2t} - 2 \log l_{x+t} + \log l_x}.$$

From this equation c is determined and then in succession $\log g$, $\log s$, and $\log k$.

52. For example, according to the Makehamized American Experience Table, we have,

$$\log k + 20 \log s + c^{20} \log g = \log l_{20} = 4.96668$$

$$\log k + 40 \log s + c^{40} \log g = \log l_{40} = 4.89286$$

$$\log k + 60 \log s + c^{60} \log g = \log l_{60} = 4.76202$$

$$\log k + 80 \log s + c^{80} \log g = \log l_{80} = 4.16122$$

$$20 \log s + c^{20}(c^{20} - 1) \log g = -.07382$$

$$20 \log s + c^{40}(c^{20} - 1) \log g = -.13084$$

$$20 \log s + c^{60}(c^{20} - 1) \log g = -.60080$$

$$c^{20}(c^{20} - 1)^2 \log g = -.05702$$

$$c^{40}(c^{20} - 1)^2 \log g = -.46996$$

Taking logarithms

$$20 \log c + 2 \log (c^{20} - 1) + \log \log g = \bar{2}.75603n$$

$$40 \log c + 2 \log (c^{20} - 1) + \log \log g = \bar{1}.67206n$$

$$20 \log c = .91603$$

$$\log c = .0458015 \quad . \quad . \quad (a)$$

Hence

$$c^{20} = 8.242$$

$$\log (c^{20} - 1) = .85986$$

$$20 \log c + 2 \log (c^{20} - 1) = 2.63575$$

$$\log \log g = \bar{2}.75603n - 2.63575 = \bar{4}.12028n$$

$$\log g = -.00013191 \quad . \quad (b)$$

$$20 \log c + \log (c^{20} - 1) + \log \log g = \bar{2}.75603n - .85986$$
$$= \bar{3}.89617n$$
$$c^{20}(c^{20} - 1) \log g = -.00787$$

but

$$20 \log s + c^{20}(c^{20} - 1) \log g = -.07382$$
$$\therefore \quad 20 \log s = -.06595$$
$$\log s = -.0032975 \quad . \quad . \quad (c)$$
$$20 \log c + \log \log g = \bar{3}.03631n$$
$$c^{20} \log g = -.00109$$
$$20 \log s + c^{20} \log g = -.06704$$

but

$$\log k + 20 \log s + c^{20} \log g = \quad 4.96668$$
$$\therefore \quad \log k = \quad 5.03372 \quad . \quad . \quad (d)$$

53. It is interesting to compare the values thus calculated with those from which the table was constructed. The differences arise from the fact that in the mortality table the value of l_x is expressed to the nearest unit and that we have used five-figure logarithms throughout. The comparison is as follows:

	Exact Value.	Calculated Value.
$\log c$	.04579609	.045802
$\log s$	−.003296862	−.003298
$\log g$	−.00013205	−.00013191
$\log k$	5.03370116	5.03372

The value of $\log c$ in the Makehamized American Experience Table is one of the highest which it has been found necessary to use. The range of values lies between .036 and .046 with a general average close to .04.

54. Unfortunately for the widest usefulness of Makeham's formula it is not possible to evaluate the integral $\int ks^x g^{c^x} dx$ otherwise than approximately, so that it does not serve the original purpose for which a mathematical law was sought. Any mathematical law, however, gives a very smooth series

which enables formulas of approximate summation or integration to be used which greatly reduce the labor of calculating the values of complicated benefits, and Makeham's law in particular offers great advantage in the calculation of probabilities of survival involving more than one life, on account of the form of the expression for $\log {}_np_x$. This point will be discussed further in connection with those probabilities. The advantage thus secured is so great, however, that it is considered that a mortality table which is to be used for monetary calculations should be adjusted so as to conform to Makeham's law if it can be accomplished without departing too seriously from the facts upon which it is based. The general subject of adjustment or graduation will be taken up in a later chapter. Modifications have been proposed to Makeham's formula for the purpose of making it fit certain tables more closely. These modifications consist in adding terms to the expression for μ_x. One modification assumes $\mu_x = A + Hx + Bc^x$, adding the term Hx to Makeham's expression, and another takes the form $\mu_x = ma^x + nb^x$.

Both of these modifications sacrifice a considerable portion of the advantage which can be secured from the use of Makeham's formula in its original form.

55. Another formula has been proposed by Wittstein, which is intended to cover the entire range of life from infancy to extreme old age. He assumes that the values of q_x may be expressed in terms of x as follows:

$$q_x = a^{-(M-x)^n} + \frac{1}{m} a^{-(mx)^n}.$$

From the form of this expression it is evident that the first term becomes equal to unity when $x = M$ and that where a is greater than unity and n is positive the value of this term increases regularly up to that value. It appears therefore that $M+1$ must equal the limiting age in the table. Also for infant mortality we have $q_0 = a^{-M^n} + \frac{1}{m}$, showing that the

probability of death during the first year after birth is slightly greater than $\frac{1}{m}$. Also

$$\frac{dq_x}{dx}=\{n(M-x)^{n-1}a^{-(M-x)^n}-n(mx)^{n-1}a^{-(mx)^n}\}\log_e a.$$

Now it is evident that this vanishes when

$$mx=(M-x) \quad \text{or} \quad x=\frac{M}{m+1},$$

and it will be found that for all cases arising in practice this represents a minimum value of q_x. In applying this formula to mortality tables it is found that for normal mortality in temperate climates a is approximately 1.42, n is approximately .63, M is between 95 and 100 and m is not less than 6. With these values it is evident that the second part of the expression for q_x decreases rapidly as x increases and becomes negligible at about age 25, so that the first term may be taken to represent adult mortality and the second term to represent the additional mortality of infancy. This formula does not possess the practical advantages of Makeham's and consequently has not been much used in practice.

56. Still another method was adopted by Prof. Karl Pearson, who took the numbers dying at the various ages and analyzed the series into the sum of five frequency curves typical respectively of old age, middle life, youth, childhood, and infancy. The table selected was that known as the English Life Table No. 4 (males) and the expression which he deduced for $l_x\mu_x$ was as follows:

$$\begin{aligned} l_x\mu_x = {} & 15.2\left(1-\frac{x-71.5}{35}\right)^{7.7525} e^{.2215(x-71.5)}, \\ & +5.4e^{-[.05524(x-41.5)]^2} \\ & +2.6e^{-[.09092(x-22.5)]^2} \\ & +8.5(x-2)^{.3271}e^{-.3271(x-3)} \\ & +415.6(x+.75)^{-.5}e^{-.75(x+.75)} \end{aligned}$$

In the first four curves the maximum values are at ages 71.5, 41.5, 22.5, and 3 respectively, while the fifth theoretically extends below age zero, the ordinate becoming infinite at age $-.75$. The method has not, however, been applied to other tables and it is difficult to lay a firm foundation for it, because no analysis of the deaths into natural divisions by causes or otherwise has yet been made such that the totals in the various groups would conform to these frequency curves.

CHAPTER IV

PROBABILITIES INVOLVING MORE THAN ONE LIFE

57. In calculating probabilities involving more than one life it is usual to assume that the probabilities of survival or death of the various lives involved are independent of one another, so that the probability of a compound event is found by simply multiplying together the elementary probabilities of which it is composed. For example, the probability that two lives now aged x and y respectively will both be alive at the end of n years is found by multiplying the probability that the life aged x will be alive, ${}_np_x$, by the probability that the life aged y will be alive, ${}_np_y$. For the sake of brevity it is usual to write (x) for a life aged x. If then the probability that both (x) and (y) will survive n years be denoted by ${}_np_{xy}$, we have

$$ {}_np_{xy} = {}_np_x \cdot {}_np_y = \frac{l_{x+n} \cdot l_{y+n}}{l_x \cdot l_y}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Similarly, where more than two lives are involved, we have

$$ {}_np_{xyz} \ldots = {}_np_x \cdot {}_np_y \cdot {}_np_z \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

58. These probabilities of joint survival are the elementary forms to which other probabilities are usually reduced. It is interesting to investigate the form which they take when Makeham's law applies. We have

$$\log {}_np_x = \log l_{x+n} - \log l_x,$$

$$= \{\log k + (x+n) \log s + c^{x+n} \log g\} - \{\log k + x \log s + c^x \log g\}$$

$$= n \log s + c^x(c^n - 1) \log g.$$

Similarly,

$$\log {}_np_y = n \log s + c^y(c^n - 1) \log g.$$

Therefore,

$$\log {}_np_{xy} = 2n \log s + (c^x + c^y)(c^n - 1) \log g$$

Let us now take two lives of equal age w. Then we have

$$\log {}_np_{ww} = 2n \log s + 2c^w(c^n - 1) \log g.$$

If, therefore, $2c^w = c^x + c^y$, the value of ${}_np_{ww}$ will be the same as that of ${}_np_{xy}$ for all values of n. Thus in all questions relating to the joint continuance of two lives aged x and y we may substitute two lives of equal ages w. The relation $2c^w = c^x + c^y$ may be expressed in another form by dividing through by c^x when we have $2c^{w-x} = 1 + c^{y-x}$, from which it follows that the value of $w - x$ depends only on that of $y - x$ and is independent of the actual values of x and y. Similarly for any number m of lives (x), (y), (z), etc., we have

$$\log {}_np_{xyz\,\ldots\,(m)} = \Sigma \log {}_np_x = mn \log s + (c^n - 1)\ \Sigma c^x \log g,$$
$$= m\{n \log s + (c^n - 1)\ c^w \log g\} = m \log {}_np_w,$$

provided $mc^w = \Sigma c^x$, so that m lives of equal ages may be substituted for any m lives.

59. Under Gompertz's law this relation takes a simple form because the term involving $\log s$ disappears and we have

$$\log {}_np_{xyz\,\ldots\,(m)} = (c^n - 1)\ \Sigma c^x \log g = (c^n - 1)\ c^w \log g = \log {}_np_w,$$

provided $c^w = \Sigma c^x$, so that here a single life may be substituted for any number of lives. In this case, too, the addition of the same number of years to each of the ages x, y, z, etc., will add the same number of years to w. This property of Gompertz's and Makeham's laws is known as the property of uniform seniority.

60. Another way in which the principle can be applied to Makeham's law is by constructing a hypothetical mortality table such that $l'_x = ks^{mx}g^{c^x}$, so that we have

$$\log {}_np'_x = mn \log s + c^x(c^n - 1) \log g.$$

We have, therefore,

$${}_np'_w = {}_np_{xyz\,\ldots\,(m)}$$

for all values of n provided $c^w = \Sigma c^x$ as in Gompertz's law. It would thus be necessary to construct a special table for each value of m, but once constructed it would apply to all combinations of m lives.

61. Hardy's modification of Makeham's law may be written

$$l_x = ks^x r^{x^2} g^{c^x},$$

or

$$\log l_x = \log k + x \log s + x^2 \log r + c^x \log g,$$

$$\begin{aligned}\log {}_n p_x = \log l_{x+n} - \log l_x &= n \log s + (2nx + n^2) \log r + c^x(c^n - 1) \log g,\\ &= (n \log s + n^2 \log r) + 2nx \log r\\ &\quad + c^x(c^n - 1) \log g.\end{aligned}$$

$$\begin{aligned}\log {}_n p_{xyz \ldots (m)} &= m(n \log s + n^2 \log r) + 2n\, \Sigma x \log r\\ &\quad + \Sigma c^x(c^n - 1) \log g,\end{aligned}$$

$$m \log {}_n p_w = m(n \log s + n^2 \log r) + 2nmw \log r + mc^w(c^n - 1) \log g.$$

The first term in each of these expressions is the same and the last terms can be made equal by putting, as in Makeham's law, $mc^w = \Sigma c^x$. This will leave an outstanding difference in the second term of $2n \log r(\Sigma x - mw)$. Since an addition of t years to each of the ages will add the same number of years to w and leave this expression unchanged, it follows that its value depends only on the differences of the ages. Since the variable n enters in the same way into this expression as into $mn \log s$, we may consider it as a modification to be applied to the value of s. In fact, if we put

$$\log s' = \log s + 2 \log r\left(\frac{\Sigma x}{m} - w\right),$$

and

$$\log {}_n p'_w = (n \log s' + n^2 \log r) + 2nw \log r + (c^n - 1)\, c^w \log g$$

then we have

$$\log {}_n p_{xyz \ldots (m)} = m \log {}_n p'_w. \tag{3}$$

It will be seen, however, that this modified value of s depends not only on m, but also on the differences of the ages, so that the complications are considerably increased.

62. If we assume

$$l_x = kr^{a^x}s^{b^x},$$

or

$$\log l_x = \log k + a^x \log r + b^x \log s,$$

we have,

$$\log {}_np_x = (a^n - 1)\, a^x \log r + (b^n - 1)\, b^x \log s.$$

If, therefore, w is determined by the equation

$$a^w/b^w = \Sigma a^x/\Sigma b^x,$$

and l is determined so that

$$l = \Sigma a^x/a^w = \Sigma b^x/b^w,$$

then we have

$$\log {}_np_{xyz\ldots(m)} = (a^n - 1)\, la^w \log r + (b^n - 1)\, lb^w \log s = l \log {}_np_w.$$

From this it follows that for the m lives (x), (y), (z), etc., we may substitute l lives of equal ages w. The difficulty is that l is not usually integral and it would, in practice, be found necessary to determine any required value by a double interpolation because w also is usually not integral.

63. We have seen that for a single life (x) we have the relation $e_x = \Sigma_n p_x$. Similarly out of a large number N of groups of m lives aged respectively x, y, z, etc., we find $N {}_1p_{xyz\ldots(m)}$ complete the first year, $N {}_2p_{xyz\ldots(m)}$ complete the second, and so on, so that if we denote by $e_{xyz\ldots(m)}$ the average number of years completed during the joint continuance of the m lives, we have

$$e_{xyz\ldots(m)} = \Sigma_n p_{xyz\ldots(m)}, \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Consequently where any expression occurs involving a summation with respect to n of the probabilities of joint survival, we may substitute a joint expectation.

64. Heretofore we have dealt with the probability that all of the lives involved shall survive. Similar reasoning will, however, show that the probability that every one of the m lives will be dead at the end of n years is obtained by multiplying together individual probabilities of death. This probability is expressed by

$$\left.\right|_n q_{\overline{xyz\ldots(m)}} \quad \text{or} \quad 1 - {}_np_{\overline{xyz\ldots(m)}}.$$

We have therefore,

$$|_{n}q_{\overline{xyz\,\ldots\,(m)}} = |_{n}q_x \cdot |_{n}q_y \cdot |_{n}q_z \ldots = (1-{}_{n}p_x)(1-{}_{n}p_y)(1-{}_{n}p_z). \quad (5)$$

In this symbol the bar over the letters denoting the ages of the lives involved signifies that the last survivor of the lives is in question, the probability designated by $|_{n}q_{\overline{xyz\,\ldots\,(m)}}$ being that the last survivor of the m lives shall have died before the end of the nth year.

65. The complementary probability is

$${}_{n}p_{\overline{xyz\,\ldots\,(m)}} = 1-(1-{}_{n}p_x)(1-{}_{n}p_y)\ \ldots,$$

and is evidently the probability that at least one of the lives will survive n years. By expanding the product and reducing the equation may be written as follows:

$${}_{n}p_{\overline{xyz\,\ldots\,(m)}} = \Sigma\,{}_{n}p_x - \Sigma\,{}_{n}p_{xy} + \Sigma\,{}_{n}p_{xyz} - \text{etc.} \quad . \quad . \quad (6)$$

In this equation the summation extends over all probabilities similar to the one under the Σ, that is, involving the same number of lives.

66. Let us now investigate the probability that exactly r out of the m lives will be alive at the end of n years. This probability is designated by ${}_{n}p_{\frac{[r]}{xyz\,\ldots\,(m)}}$. The probability that r particular lives, (x), (y), etc., are alive and the remaining $(m-r)$ lives (z), (w), etc., all dead is evidently

$${}_{n}p_{xy\,\ldots\,(r)} \cdot |_{n}q_{\overline{zw\,\ldots\,(m-r)}} \quad \text{or} \quad {}_{n}p_{xy\,\ldots\,(r)}(1-{}_{n}p_z)(1-{}_{n}p_w)\ \ldots$$

and the total probability sought is the sum of these probabilities for all the combinations r at a time of the m lives, or

$${}_{n}p_{\frac{[r]}{xyz\,\ldots\,(m)}} = \Sigma\,{}_{n}p_{xy\,\ldots\,(r)}(1-{}_{n}p_z)(1-{}_{n}p_w) \quad . \quad . \quad . \quad (7)$$

From the form of the expression it is evident that it may be expanded in a series of probabilities of joint survival involving from r up to m lives; also that each probability involving more than r lives will appear more than once in the expression because it will appear once for each combination r at a time of the lives involved in it; also that the sign of any probability in-

volving $r+t$ lives is positive or negative according as t is even or odd. Thus we have

$$
\begin{aligned}
{}_{n}p_{\overline{xyz\,.\,.\,.\,(m)}}^{[r]} &= \Sigma_n p_{xy\,.\,.\,.\,(r)} - {}_{r+1}C_r \Sigma_n p_{xy\,.\,.\,.\,(r+1)} + {}_{r+2}C_r \Sigma_n p_{xy\,.\,.\,.\,(r+2)} \\
&\qquad - \text{etc.} \\
&= \Sigma_n p_{xy\,.\,.\,.\,(r)} - (r+1)\Sigma_n p_{xy\,.\,.\,.\,(r+1)} \\
&\qquad + \frac{(r+1)(r+2)}{1\cdot 2}\Sigma_n p_{xy\,.\,.\,.\,(r+2)} - \text{etc.} \quad . \quad (8)
\end{aligned}
$$

67. This may be verified by supposing all the ages x, y, z, etc., to be equal, in which case $\Sigma_n p_{xyz\,.\,.\,.\,(r+t)}$ becomes equal to ${}_{m}C_{r+t}\,{}_{n}p_x^{\,r+t}$, because ${}_{m}C_{r+t}$ is the number of terms included in the summation and each term becomes equal to ${}_{n}p_x^{\,r+t}$. The whole expression therefore reduces to

$$
\begin{aligned}
&{}_{m}C_r\,{}_{n}p_x^{\,r} - {}_{r+1}C_r \cdot {}_{m}C_{r+1} \cdot {}_{n}p_x^{\,r+1} + {}_{r+2}C_r \cdot {}_{m}C_{r+2}\,{}_{n}p_x^{\,r+2} - \text{etc.}, \\
&= {}_{m}C_r \cdot {}_{n}p_x^{\,r} - {}_{m}C_r \cdot {}_{m-r}C_1 \cdot {}_{n}p_x^{\,r+1} + {}_{m}C_r \cdot {}_{m-r}C_2 \cdot {}_{n}p_x^{\,r+2} - \text{etc.} \\
&= {}_{m}C_r \cdot {}_{n}p_x^{\,r}(1 - {}_{n}p_x)^{m-r}.
\end{aligned}
$$

This is evidently the proper expression for the probability in question, because the probability that any particular r of the m lives aged x are all alive and the remaining $(m-r)$ all dead is ${}_{n}p_x^{\,r}(1-{}_{n}p_x)^{m-r}$, and there are ${}_{m}C_r$ different groups of r lives included among the m.

68. A very convenient symbolic notation is sometimes used to condense the form of Eq. (8) by substituting Z^t for $\Sigma_n p_{xyz\,.\,.\,.\,(t)}$, when the equation takes the form

$$
\begin{aligned}
{}_{n}p_{\overline{xyz\,.\,.\,.\,(m)}}^{[r]} &= Z^r - (r+1)Z^{r+1} + \frac{(r+1)(r+2)}{1\cdot 2}Z^{r+2} \\
&\qquad - \frac{(r+1)(r+2)(r+3)}{1\cdot 2\cdot 3}Z^{r+3} + \text{etc.}, \\
&= Z^r(1+Z)^{-(r+1)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)
\end{aligned}
$$

In this connection it is to be remembered that the expression is purely symbolic and that no operations can be performed upon it which in any way disturb the meaning of Z^t.

69. Let us now investigate the probability that at least r out of m lives will survive n years. This is denoted by ${}_np_{\frac{r}{xyz\ldots(m)}}$ and it is evident that we have

$$ {}_np_{\frac{r}{xyz\ldots(m)}} = {}_np_{\frac{[r]}{xyz\ldots(m)}} + {}_np_{\frac{[r+1]}{xyz\ldots(m)}} + \ldots\ {}_np_{xyz\ldots(m)}. \qquad (10) $$

From this we see that the expression may be written in the form

$$ {}_np_{\frac{r}{xyz\ldots(m)}} = \Sigma\,{}_np_{xyz\ldots(r)} + a_1\Sigma\,{}_np_{xyz\ldots(r+1)} + a_2\Sigma\,{}_np_{xyz\ldots(r+2)} + \text{etc.} $$

where a_1, a_2, etc., remain to be determined. But from Eqs. (8) and (10) we have

$$ \begin{aligned} a_t &= 1 - {}_{r+t}C_1 + {}_{r+t}C_2 - {}_{r+t}C_3 + \ldots + (-1)^t\,{}_{r+t}C_t, \\ &= 1 - (1 + {}_{r+t-1}C_1) + ({}_{r+t-1}C_1 + {}_{r+t-1}C_2) - ({}_{r+t-1}C_2 + {}_{r+t-1}C_3) \\ &\qquad\qquad + \ldots + (-1)^t({}_{r+t-1}C_{t-1} + {}_{r+t-1}C_t) \\ &= (-1)^t\,{}_{r+t-1}C_t. \end{aligned} $$

Therefore, we have

$$ \begin{aligned} {}_np_{\frac{r}{xyz\ldots(m)}} &= \Sigma\,{}_np_{xyz\ldots(r)} - r\Sigma\,{}_np_{xyz\ldots(r+1)} + \frac{r(r+1)}{2}\Sigma\,{}_np_{xyz\ldots(r+2)} \\ &\qquad\qquad - \text{etc.}, \\ &= Z^r - rZ^{r+1} + \frac{r(r+1)}{2}Z^{r+2} - \frac{r(r+1)(r+2)}{6}Z^{r+3} + \text{etc.} \\ &= Z^r(1+Z)^{-r}, \qquad \ldots\ldots\ldots\ldots \qquad (11) \end{aligned} $$

where the same meaning is assigned to Z^r as before.

70. It is to be noted that although the relation is purely symbolic and the function of Z has no meaning except as expanded in ascending powers and then interpreted, we have the following relation:

$$ \begin{aligned} {}_np_{\frac{r}{xyz\ldots(m)}} &= Z^r(1+Z)^{-r} = Z^r(1+Z)^{-(r+1)}\{1 - Z(1+Z)^{-1}\}^{-1} \\ &= Z^r(1+Z)^{-(r+1)}\{1 + Z(1+Z)^{-1} + Z^2(1+Z)^{-2} + \ldots\} \\ &= Z^r(1+Z)^{-(r+1)} + Z^{r+1}(1+Z)^{-(r+2)} + Z^{r+2}(1+Z)^{-(r+3)} \\ &\qquad + \ldots \text{ etc.}, \\ &= {}_np_{\frac{[r]}{xyz\ldots(m)}} + {}_np_{\frac{[r+1]}{xyz\ldots(m)}} + {}_np_{\frac{[r+2]}{xyz\ldots(m)}} + \text{etc.}, \end{aligned} $$

as in Eq. (10).

71. Also if we have m lives aged respectively x, y, z, etc., the expected number of survivors at the end of n years is $\Sigma r_n p_{\overset{[r]}{xyz\ldots(m)}}$, where the summation extends over all values of r from unity to m. Expressed symbolically this becomes, from Eq. (9),

$$\begin{aligned} &Z(1+Z)^{-2}+2Z^2(1+z)^{-3}+3Z^3(1+Z)^{-4}+\text{etc.},\\ &\qquad =Z(1+Z)^{-2}\{1+2Z(1+Z)^{-1}+3Z^2(1+Z)^{-2}+\text{etc.}\},\\ &\qquad =Z(1+Z)^{-2}\{1-Z(1+Z)^{-1}\}^{-2},\\ &\qquad =Z=\Sigma_n p_x. \qquad (12) \end{aligned}$$

This may be also verified by reasoning similar to that by which Eq. (11) was deduced. We thus see that the expected number of survivors out of any group of lives is found by adding together the individual probabilities of survival.

72. Another class of probabilities involving more than one life relates to the order in which the deaths occur. The probability that (x) will die in the nth year from the present time is

$$\frac{d_{x+n-1}}{l_x}=\frac{l_{x+n-1}-l_{x+n}}{l_x}={}_{n-1}p_x-{}_np_x=\frac{{}_np_{x-1}}{p_{x-1}}-{}_np_x,$$

and the probability that (y) will be alive at the end of the nth year is ${}_np_y$. The probability therefore that (x) will die in the nth year and (y) will be alive at the end of that year is

$$\left(\frac{{}_np_{x-1}}{p_{x-1}}-{}_np_x\right){}_np_y=\frac{{}_np_{\overline{x-1}:y}}{p_{x-1}}-{}_np_{xy}.$$

Summing this function for all values of n from unity up, we get the total probability that (y) will be alive at the end of the year in which the death of (x) occurs This sum is

$$\frac{1}{p_{x-1}}\Sigma_{n=1}^{n=\infty}{}_np_{\overline{x-1}:y}-\Sigma_{n=1}^{n=\infty}{}_np_{xy}=\frac{1}{p_{x-1}}e_{\overline{x-1}:y}-e_{xy}.$$

Similarly, the probability that (x) will be alive at the end of the year in which the death of (y) occurs is $\frac{1}{p_{y-1}}e_{x:\overline{y-1}}-e_{xy}$,

and the probability that both deaths will occur in the same year is the complement of the sum of these probabilities and is therefore,

$$1+2e_{xy}-\frac{1}{p_{x-1}}e_{\overline{x-1}:y}-\frac{1}{p_{y-1}}e_{x:\overline{y-1}}.$$

It may be assumed that where both deaths occur in the same year the chances are even that the death (x) will occur before that of (y). The total chance therefore that (x) will die before (y) denoted by Q^1_{xy} is

$$Q^1_{xy}=\left\{\frac{1}{p_{x-1}}e_{\overline{x-1}:y}-e_{xy}\right\}+\frac{1}{2}\left\{1+2e_{xy}-\frac{1}{p_{x-1}}e_{\overline{x-1}:y}-\frac{1}{p_{y-1}}e_{x:\overline{y-1}}\right\}$$

$$=\frac{1}{2}\left\{1+\frac{1}{p_{x-1}}e_{\overline{x-1}:y}-\frac{1}{p_{y-1}}e_{x:\overline{y-1}}\right\} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (13)$$

73. The same probability may be otherwise expressed in terms of the infinitesimal calculus by indefinitely reducing the intervals considered. The probability that the death of (x) will occur in the interval of time between t and $t+dt$ will be $-\frac{dl_{x+t}}{dt}dt/l_x$, and the probability that (y) will be alive at that time is $\frac{l_{y+t}}{l_y}$. The total probability therefore that (y) will be alive when the death of (x) occurs is

$$Q^1_{xy}=-\frac{1}{l_xl_y}\int_0^\infty l_{y+t}\frac{dl_{x+t}}{dt}dt,$$

$$=-\frac{1}{l_xl_y}\int_0^\infty l_{y+t}\frac{dl_{x+t}}{dx}dt,$$

$$=-\frac{1}{l_xl_y}\cdot\frac{d}{dx}\int_0^\infty l_{x+t}\,l_{y+t}dt,$$

$$=-\frac{1}{l_xl_y}\cdot\frac{d}{dx}(l_xl_y\mathring{e}_{xy}),$$

$$=-\frac{1}{l_x}\frac{d}{dx}(l_x\mathring{e}_{xy}),$$

$$=-\frac{1}{l_x}\left(\mathring{e}_{xy}\frac{dl_x}{dx}+l_x\frac{d\mathring{e}_{xy}}{dx}\right),$$

$$=\frac{1}{l_x}\left(\mathring{e}_{xy}l_x\mu_x - l_x\frac{de_{xy}}{dx}\right),$$

$$=\mu_x\mathring{e}_{xy} - \frac{d\mathring{e}_{xy}}{dx},$$

$$=\mu_x\mathring{e}_{xy} + \tfrac{1}{2}(\mathring{e}_{x-1:y} - \mathring{e}_{x+1:y}) \text{ approximately.} \quad (14)$$

74. Similarly, where m lives, (x), (y), (z), etc., are involved, the probability that (x) will die first is

$$Q^1_{xyz\ldots(m)} = -\frac{1}{l_xl_y\ldots}\int_0^\infty \frac{dl_{x+t}}{dt}\cdot l_{y+t}\cdot l_{z+t}\ldots dt,$$

$$=\mu_x\mathring{e}_{xyz\ldots(m)} + \tfrac{1}{2}(\mathring{e}_{x-1:yz\ldots(m)} - \mathring{e}_{x+1:yz\ldots(m)}). \quad (15)$$

75. Also from the fact that the probability that (x) will die rth in order out of the group of m lives may also be stated as the probability that, when (x) dies, there will be exactly $m-r$ survivors of the $m-1$ lives other than (x), we may express this probability by the use of Eq. (9) in terms of probabilities of dying first. In fact, if we denote by Y^t the sum of the values of $Q^1_{xyz\ldots}$ for (x) along with all groups t at a time of the $(m-1)$ lives (y), (z), etc., we have

$$Q^r_{xyz\ldots(m)} = \int_0^\infty \mu_{x+t}\ {}_tp_x Z^{m-r}(1+Z)^{-(m-r+1)}dt,$$

where the summation included in Z covers only the $m-1$ lives (y), (z), etc.

But $\int_0^\infty \mu_{x+t}\ {}_tp_x Z^l dt = Y^l$ for all values of l. Therefore expanding, integrating, and condensing, we have

$$Q^r_{xyz\ldots(m)} = Y^{m-r}(1+Y)^{-(m-r+1)}. \quad (16)$$

For $m=3$ we have

$$Q^1_{xyz} = \mu_x\mathring{e}_{xyz} + \tfrac{1}{2}(\mathring{e}_{x-1:yz} - \mathring{e}_{x+1:yz}), \quad (17)$$

$$Q^2_{xyz} = Y(1+Y)^{-2} = Y - 2Y^2,$$

$$= Q^1_{xy} + Q^1_{xz} - 2Q^1_{xyz}, \quad (18)$$

$$Q^3_{xyz} = (1+Y)^{-1} = 1 - Y + Y^2,$$

$$= 1 - Q^1_{xy} - Q^1_{xz} + Q^1_{xyz}. \quad (19)$$

76. Where Makeham's law is assumed to hold, the probability $\overset{1}{Q}_{xy}$ takes a special form. We have generally

$$\overset{1}{Q}_{xy} = -\frac{1}{l_x l_y}\int_0^{\infty} l_{y+t}\frac{dl_{x+t}}{dt}dt = \frac{1}{l_x l_y}\int_0^{\infty} l_{y+t}l_{x+t}\mu_{x+t}dt,$$

$$= \int_0^{\infty} \mu_{x+t}\,{}_tp_{xy}dt,$$

$$= \int_0^{\infty} (A + Bc^{x+t})\,{}_tp_{xy}dt,$$

$$= A\int_0^{\infty} {}_tp_{xy}dt + Bc^x\int_0 c^t\,{}_tp_{xy}dt,$$

$$= A\mathring{e}_{xy} + Bc^x\int_0^{\infty} c^t\,{}_tp_{xy}dt.$$

Similarly,

$$Q_{x\overset{1}{y}} = A\mathring{e}_{xy} + Bc^y\int_0^{\infty} c^t\,{}_tp_{xy}dt.$$

Therefore,

$$1 = \overset{1}{Q}_{xy} + Q_{x\overset{1}{y}} = 2A\mathring{e}_{xy} + B(c^x + c^y)\int_0^{\infty} c^t\,{}_tp_{xy}dt,$$

or

$$B\int_0^{\infty} c^t\,{}_tp_{xy}dt = \frac{1}{c^x + c^y}(1 - 2A\mathring{e}_{xy}).$$

Substituting this in the expression for $\overset{1}{Q}_{xy}$, we have

$$\overset{1}{Q}_{xy} = A\mathring{e}_{xy} + \frac{c^x}{c^x + c^y}(1 - 2A\mathring{e}_{xy}),$$

$$= \frac{c^x}{c^x + c^y} - A\frac{c^x - c^y}{c^x + c^y}\mathring{e}_{xy},$$

$$= \frac{1}{1 + c^{y-x}} - A\frac{1 - c^{y-x}}{1 + c^{y-x}}\mathring{e}_{xy}. \quad . \quad . \quad . \quad . \quad (20)$$

77. For Gompertz's law we have $A = 0$, and this equation takes the form

$$\overset{1}{Q}_{xy} = \frac{1}{1 + c^{y-x}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

We thus see that according to this law the probability of survivorship depends only on the difference of ages.

CHAPTER V

STATISTICAL APPLICATIONS

78. SUPPOSE that in a certain country there is neither emigration nor immigration and that the number of births occurring each year is uniform and equal to l_0. Then it is evident that on the assumption that the law of mortality remains unchanged the number of inhabitants attaining age x in each year is l_x, being the survivors out of the l_0 who were born x years before. Therefore, at any moment the number in the existing population whose age is between x and $x+dx$ will be $l_x dx$, being the survivors out of the $l_0 dx$ who were born in the interval dx years x years earlier. The total population at age x last birthday or between ages x and $x+1$ is, therefore,

$$\int_x^{x+1} l_x dx = \int_0^1 l_{x+t} dt.$$

If now we assume that the deaths are evenly distributed, so that

$$l_{x+t} = l_x - td_x,$$

we have for the population at age x last birthday denoted by L_x the following

$$L_x = \int_0^1 (l_x - td_x)dt = l_x - \tfrac{1}{2}d_x = \tfrac{1}{2}(l_x + l_{x+1}). \quad . \quad . \quad (1)$$

Also the deaths occurring per annum between the ages x and $x+dx$ will be $l_x\mu_x dx = -\dfrac{dl_x}{dx}dx$. Therefore, the total deaths per annum at age x last birthday or between ages x and $x+1$ will be $\displaystyle\int_x^{x+1} -\frac{dl_x}{dx}dx = l_x - l_{x+1} = d_x$. Summing this for all ages x and over we see that the total deaths for those ages is l_x, so that the aggregate number of deaths per annum at all ages

will be l_0, which is also the number of births. The population is therefore constant in total number and also in age composition.

79. The total population at age x and over will be $\Sigma_x^w L_x$, which is usually denoted by T_x, so that we have

$$T_x = \Sigma_x^w L_x = \Sigma_x^w \tfrac{1}{2}(l_x + l_{x+1}) = \tfrac{1}{2}l_x + \Sigma_{x+1}^w l_x = l_x \mathring{e}_x. \quad . \quad (2)$$

The total population at all ages will be

$$T_0 = l_0 \mathring{e}_0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The general average death rate is obtained by dividing the total number of deaths per annum by the total population, its value is therefore equal to $l_0/T_0 = l_0/l_0\mathring{e}_0 = 1/\mathring{e}_0$ or the reciprocal of the complete expectation of life at birth. Similarly, the average death rate at ages x and over will be

$$l_x/T_x = l_x/l_x\mathring{e}_x = 1/\mathring{e}_x.$$

The total number of deaths per annum between ages x and $x+n$ will be $l_x - l_{x+n}$, and the total population at the same ages will be $T_x - T_{x+n}$, therefore, the average death rate between those ages will be $\dfrac{l_x - l_{x+n}}{T_x - T_{x+n}}$. When n is equal to 1, this becomes the central death rate for age x last birthday and is denoted by m_x, so that we have

$$m_x = \frac{l_x - l_{x+1}}{T_x - T_{x+1}} = \frac{d_x}{L_x} = \frac{d_x}{\frac{1}{2}(l_x + l_{x+1})} = \frac{d_x}{l_x - \frac{1}{2}d_x}. \quad . \quad . \quad (4)$$

If we divide both numerator and denominator by l_x, we have m_x in terms of q_x, as follows:

$$m_x = \frac{q_x}{1 - \frac{1}{2}q_x} \quad \text{or} \quad \frac{1}{m_x} = \frac{1}{q_x} - \tfrac{1}{2}. \quad . \quad . \quad . \quad . \quad (5)$$

80. Assuming uniform distribution of deaths within each year of age, the sum of the ages at death of all those dying in a year at ages x and over will be

$$\begin{aligned}\Sigma_x^w (x + \tfrac{1}{2}) d_x &= \Sigma_x^w (x + \tfrac{1}{2})(l_x - l_{x+1}), \\ &= (x + \tfrac{1}{2}) l_x + \Sigma_{x+1}^w l_x = x l_x + \Sigma_x^w L_x = x l_x + T_x, \\ &= (x + \mathring{e}_x) l_x.\end{aligned}$$

Since the total number of deaths at those ages is l_x it follows that the average age is $x+\mathring{e}_x$. Putting x equal to zero, we see that the general average age at death is $\mathring{e}_0$.

The aggregate of the ages at death of those dying between ages x and $x+n$ is evidently

$$(xl_x+T_x)-\{(x+n)l_{x+n}+T_{x+n}\}=(T_x-T_{x+n})+\{xl_x-(x+n)l_{x+n}\}$$

and the number of deaths is l_x-l_{x+n}, so that the average age is

$$\frac{(T_x-T_{x+n})+xl_x-(x+n)l_{x+n}}{l_x-l_{x+n}}=x+\frac{T_x-T_{x+n}-nl_{x+n}}{l_x-l_{x+n}}.$$

81. We have seen that the population at age x last birthday is L_x, and the total population is T_0, so that the proportion of the total population at age x last birthday is L_x/T_0 and the proportion between ages x and $x+n$ is $(T_x-T_{x+n})/T_0$. Suppose, for example, that all young men are required to serve in the army from age 18 to 21, then, assuming a stationary population, the proportion of the total male population so serving will be $(T_{18}-T_{21})/T_0$.

82. We have hitherto assumed that we are dealing with a stationary population. A consideration, however, of the question leads to the conclusion that such a condition never exists, but that, owing to various disturbing factors, the percentages of the total population at the various ages will not be exactly the same as in the assumed stationary population derived from the mortality table representing the actual death rates experienced. It is evident that, if for any reason we have in one community more than the normal percentage of the population at those ages where the death rate is low and in another community less than the normal percentage at those ages, then, even though the death rate at every individual age might be the same in the two communities, the general average death rate in the first will be less than in the second. It cannot be assumed, therefore, that a higher average death rate necessarily means a more unfavorable mortality experience. A correction must first be made for the difference in age distribution.

83. One method of making this correction is to construct the mortality tables representing the observed death rates, analyzed by ages, in the two communities and calculate from such tables the complete expectation of life at birth. From the fact that in a stationary population the general average death rate is the reciprocal of this expectation it is readily seen that this amounts, in effect, to substituting for each actual population the stationary population corresponding to its actual mortality. It is readily seen that this method may be applied to the death rates for ages above any assigned age, or within given limits, by constructing the corresponding portion of the mortality table and calculating the average death rate in the stationary population. For example, it might be desired to compare the mortality in two communities for ages 15 and over or for ages 15 to 64 last birthday inclusive. The corrected death rate for the former would be $1/\mathring{e}_{15}$, and for the latter $\dfrac{l_{15}-l_{65}}{T_{15}-T_{65}}=\dfrac{\Sigma_{15}^{64}d_x}{\Sigma_{15}^{64}L_x}$. The labor of constructing a mortality table is, however, considerable and other methods of correction are usually followed.

84. Although the stationary population is largely of theoretical interest the notation derived from it is useful with certain modifications in connection with actual population statistics. For this purpose θ_x represents the deaths between age x and age $x+1$, and $\lambda_x=\theta_x+\theta_{x+1}+\text{etc.}$, is the total number of deaths at age x and over, but is not equal to the number attaining age x. For the population between ages x and $x+1$ the symbol L_x is retained. The symbol T_x is also used to denote the total population at ages x and over, so that we have

$$T_x=L_x+L_{x+1}+\text{etc.},$$

as before. The general average death rate is then λ_0/T_0, but is not equal to $1/\mathring{e}_0$ except for a stationary population. Similarly for the average death rate at age x and over we have λ_x/T_x but not $1/\mathring{e}_x$, and for ages between x and $x+n$ we have $\lambda_x-\lambda_{x+n}/T_x-T_{x+n}$.

85. One method of correcting the death rates of different communities is to analyze each into certain age groups, usually quinquennial up to age 15, then decennial up to age 85, with a final group for ages 85 or more last birthday, the average death rate for each group being used. These death rates are then applied to a standard proportionate distribution of the population into these age groups. One standard which has been used is the age distribution of the population of England and Wales according to the Census of 1801. The general average death rate for the standard population on the basis of the observed group rates for each community is thus calculated and this is considered as the corrected death rate for the community. In this way all communities entering into the comparison are placed on the same footing with respect to age distribution. The same method may be extended to cover varying proportions of the two sexes by analyzing the statistics for the different communities and also the standard population in this way. It may, in fact, be extended to cover any factor, such as occupations, considered as having an important bearing on the mortality to be expected and for which the necessary data can be obtained.

86. Another method of comparison is to use a standard scale of death rates for the different groups into which the actual populations are analyzed. The actual population in each group is then multiplied by the standard death rate and the expected deaths according to the standard are thus calculated. The total of the actual deaths in the community is then expressed as a percentage of the expected and these percentages for the different communities are compared.

87. These two methods have been described as applying to a whole community, but it is evident that they apply also to a part, such as those aged x and over, or those whose ages lie between x and $x+n$, or those who are engaged in a certain occupation. In fact, what may be considered as mortality index numbers for various occupations have been formed from the census and death returns in England. A standard population is taken, analyzed into the five decennial age groups

between 15 and 65, the aggregate population being such that the expected deaths according to the general average death rates for occupied males in the various age groups will total up to 1000. The actual death rates for the various age groups in each occupation are then applied to this standard population and the resulting total of expected deaths gives a number whose ratio to 1000 measures the general mortality of the occupation. This is in effect the standard population method above described with the addition that instead of recording the corrected average death rate we record its ratio to an average death rate based on the same standard population combined with standard group death rates.

88. The standard population method is the one most used for the comparison of general population mortality statistics, while the standard death rate method is most used in connection with the mortality of insured lives. In connection with such insurance statistics three modifications are made. The first is that the actual experience is usually analyzed into individual years of age and sometimes also into years elapsed since medical examination. The second is that the rate of mortality or probability of dying within one year is usually used instead of the death rate or average force of mortality, and that along with it the exposed to risk of death, which is discussed under the head of construction of mortality tables, must be used instead of the population. The third is that amounts insured or amounts at risk are frequently taken into account instead of lives, so that we compare actual losses with expected losses rather than actual deaths with expected deaths.

89. In this chapter it has been assumed that the period covered by the statistics is one year. Where a period other than one year is dealt with, we must take the average deaths per annum, and in any event whether for a period of exactly one year or otherwise the average population during the period must be taken. The ratio will, of course, be the same if both of these are multiplied by the period, so that we have on the one hand the total deaths and on the other the aggregate number of years of life during the period.

CHAPTER VI

CONSTRUCTION OF MORTALITY TABLES

90. In the second chapter it was shown that in any mortality table we have the relation $l_{x+1}=l_x p_x$ for all values of x and that consequently if we have a complete table of the values of p_x we can, by starting at the initial age and working forward progressively, construct a complete mortality table. A little consideration also shows us that there is an insuperable practical difficulty in the way of constructing the l_x column of a mortality table by taking a large group of lives of a given age and following them throughout the balance of their lives, observing the number surviving to each age. This difficulty arises not only from the length of time that would necessarily be consumed in waiting for the last one to die, but also from the fact that out of any large number some are certain to pass out of the knowledge of the observers and from the moment that any do so disappear the further observations are nullified by our ignorance of the time of their death. A correction is therefore necessary and this correction can be most conveniently applied by a method which also obviates the necessity of waiting until some particular group of lives selected at a young age have all died. This method is to use the relation already quoted and to determine separately the values of p_x for each year of age. By this method the observations do not necessarily extend over a longer period than one year, although a longer period is usually taken in order to eliminate the effect of special conditions. In that event the observations at different times for the same year of age are combined.

91. The observations are not, in fact, made directly on the value of p_x, but rather on that of m_x determined by the relation

$$m_x=\theta_x/L_x, \qquad \qquad (1)$$

where θ_x represents the deaths observed at age x, last birthday, and L_x is the corresponding population.

But we have in terms of the mortality table

$$\frac{m_x}{2}=\frac{d_x}{l_x+l_{x+1}}=\frac{l_x-l_{x+1}}{l_x+l_{x+1}}=\frac{1-p_x}{1+p_x},$$

from which we have

$$p_x=\frac{2-m_x}{2+m_x}, \qquad \text{. . . . . (2)}$$

and

$$q_x=1-p_x=\frac{2m_x}{2+m_x}=\frac{\theta_x}{L_x+\frac{1}{2}\theta_x}. \qquad \text{. . . . . (3)}$$

92. In connection with population statistics it has been usual to calculate m_x from the data and then to pass to q_x and p_x. In connection with observations on insured lives, on the other hand, the practice has been to determine the value of $L_x+\frac{1}{2}\theta_x$ denoted by E_x for each age and so to proceed directly to q_x by the equation $q_x=\theta_x/E_x$. The problem therefore reduces to the determination of the values of θ_x for each value of x, and of the corresponding values of L_x or E_x. The methods followed vary, of course, with the form in which the facts are presented, and the conditions in connection with general population statistics differ so much from those in connection with insured lives that it is well to take up the two cases separately.

93. In the case of general population statistics the information regarding the deaths is usually derived from the registration returns and it is a necessary condition, for their use in the determination of death rates, that the registration should include all the deaths coming within the scope of the investigation. It is evident that, to the extent that the returns are incomplete, the numerator of the fraction determining the death rate is understated and consequently the death rate itself is also understated. For this reason the statistics can be used of only those countries, states or municipalities in which the laws and their enforcement are such as to secure

substantial accuracy in the death returns. In the United States those states and parts of states which, in the opinion of the Federal Census Bureau, comply with this requirement constitute the registration district. The area included in this district is extended from time to time as the registration becomes more complete. Particulars of the deaths in the various component parts of the registration area are published annually by the Census Bureau.

94. For information regarding the population corresponding to the deaths reported we must depend upon the census results. As a census is made only periodically, some means must be devised of passing from these figures at periodical intervals to the average population by age groups during the interval covered by the observed deaths. The census returns and the death returns are also frequently given only for groups of ages, and we have therefore an additional problem to solve, namely, that of passing from age groups to individual years of age.

95. Let us take first the problem of finding the average population in a certain age group during a specified period. For the sake of simplicity we will first suppose that the total population analyzed by age groups is known for the beginning and end of the period and that the whole period may be considered for this purpose as a unit of time. Let the total population at the beginning be P_0 and at the end P_1, also let the population in any particular age group at the beginning be aP_0 and at the end $(a+b)P_1$. Then, evidently, the sum of the values of a for all age groups must be equal to unity and this is also true for the values of $(a+b)$, so that the sum of the values of b must be zero. Also suppose the ratio of increase of the total population during the period is r, so that we have $P_1=rP_0$. Then it is assumed that at any time t during the interval the population in the age group is $P_0(a+bt)r^t$, the sum of the values of which for all age groups is evidently P_0r^t. In other words the total population is supposed to vary in geometrical progression, while the percentage of that total in the particular age group is supposed to vary in arithmetical

progression. On these assumptions the average population during the period is

$$P_0\int_0^1(a+bt)r^t dt=aP_0\int_0^1 r^t dt+bP_0\int_0^1 tr^t dt,$$

$$=aP_0\frac{r-1}{\log_e r}+bP_0\left\{\frac{r}{\log_e r}-\frac{r-1}{(\log_e r)^2}\right\},$$

$$=P_0\frac{r-1}{\log_e r}\left\{a+b\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)\right\} \quad . \quad . \quad (4)$$

96. It is evident that if the period covered by the observations were the interval between two censuses, the census returns would give directly the values of P_0, P_1, a and b. But the dates upon which the census is taken do not usually coincide with the limits of the period of observations. Suppose, therefore, that we have the results of two censuses taken at the times t_1 and t_2 counting from the beginning of the period and that the corresponding total populations are P_3 and P_4, also that the populations in the age group are AP_3 and BP_4. Then, according to the assumptions already made, we have

$$P_3=r^{t_1}P_0 \quad \text{or} \quad \log P_3=\log P_0+t_1\log r,$$

$$P_4=r^{t_2}P_0 \quad \text{or} \quad \log P_4=\log P_0+t_2\log r,$$

$$(t_2-t_1)\log r=\log P_4-\log P_3,$$

$$\log r=(\log P_4-\log P_3)/(t_2-t_1), \quad . \quad . \quad . \quad . \quad (5)$$

$$\log P_0=\log P_3-t_1\log r=(t_2\log P_3-t_1\log P_4)/(t_2-t_1). \quad (6)$$

$$a+bt_1=A,$$

$$a+bt_2=B,$$

$$b(t_2-t_1)=(B-A),$$

$$b=(B-A)/(t_2-t_1), \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

$$a=A-bt_1=(t_2A-t_1B)/(t_2-t_1). \quad . \quad . \quad . \quad . \quad (8)$$

97. The expression for the average population in the age group, when we substitute in Eq. (4) these values of a and b, takes the form

$$P_0\frac{r-1}{\log_e r}\left\{\frac{t_2A-t_1B}{t_2-t_1}+\frac{B-A}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)\right\}$$

$$=P_0\frac{r-1}{\log_e r}\left\{\frac{t_2A}{t_2-t_1}-\frac{A}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)\right.$$

$$\left.-\frac{t_1B}{t_2-t_1}+\frac{B}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)\right\}$$

$$=AP_0\frac{r-1}{\log_e r}\left\{\frac{t_2}{t_2-t_1}-\frac{1}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)\right\}$$

$$+BP_0\frac{r-1}{\log_e r}\left\{\frac{1}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)-\frac{t_1}{t_2-t_1}\right\}. \quad . \quad . \quad . \quad (9)$$

Since $AP_0r^{t_1}$ and $BP_0r^{t_2}$ are the numbers shown in the two censuses for the age group in question, it follows that we obtain the average population for any age group by multiplying the numbers shown in the two censuses by

$$r^{-t_1}\frac{r-1}{\log_e r}\left\{\frac{t_2}{t_2-t_1}-\frac{1}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)\right\},$$

and

$$r^{-t_2}\frac{r-1}{\log_e r}\left\{\frac{1}{t_2-t_1}\left(\frac{r}{r-1}-\frac{1}{\log_e r}\right)-\frac{t_1}{t_2-t_1}\right\},$$

respectively, and adding together the products. These factors are the same for all age groups and may be calculated once for all.

98. The average deaths per annum may be obtained by dividing the total deaths during the period by the number of years included, or the same object can be accomplished by multiplying up the average population by the number of years to get the aggregate population or years of life corresponding to the total number of deaths. The latter is the course usually followed.

99. Having, then, the total deaths and the corresponding population by groups of ages the remaining problem is to ascertain the death rates for individual ages. An approximation which was formerly used was to divide the total deaths by the total population, and assume that this represented the force of mortality at the middle of the interval, or, in terms of the

notation explained in Chapter V, $\mu_{x+\frac{n}{2}}=\frac{\lambda_x-\lambda_{x+n}}{T_x-T_{x+n}}$. Where n is odd this gives directly the value of $m_{x+\frac{1}{2}(n-1)}$, the two functions being approximately equal and each equal to $d_{x+\frac{1}{2}(n-1)}/l_{x+\frac{1}{2}n}$. Thus $q_{x+\frac{1}{2}(n-1)}$ is obtained by Eq. (3) of Chapter VI. Where n is even, however, $x+\frac{1}{2}n$ is an integer. The value of $q_{x+\frac{1}{2}n}$ is then determined on the assumption that during the year μ_x increases in geometrical progression at the ratio r determined from the values of μ_x for the neighboring groups. We have then since $\mu_{x+t}=\frac{d \operatorname{colog}_e l_{x+t}}{dt}$

$$\operatorname{colog}_e p_{x+\frac{1}{2}n}=\int_0^1 \mu_{x+\frac{1}{2}n+t}dt=\mu_{x+\frac{1}{2}n}\int_0^1 r^t dt,$$

$$=\mu_{x+\frac{1}{2}n}\frac{r-1}{\log_e r}. \qquad \text{(10)}$$

From these values of q_x the intermediate values are found by a formula of interpolation.

100. It was always recognized that the quinquennial age group from 10 to 15 required special treatment, and it has recently been shown by Mr. Geo. King that the method understates the death rate at the older ages. This can be seen by taking any mortality table and, assuming a stationary population, comparing the values of $\frac{\lambda_x-\lambda_{x+5}}{T_x-T_{x+5}}$ with those of m_{x+2}. In view of this fact some more accurate method is desirable. Greater accuracy has been attained by distributing the total deaths and population of each age group into individual years of age.

101. In the construction of the Carlisle table this distribution was effected by a graphic method. On a base line distances were laid off consecutively representing the number of years included in the successive age groups. On these bases rectangles were constructed whose area represented the total number (of deaths or of population as the case may be) in the age group. The heights therefore represented the average number per year of age in each group. A continuous

line with continuous curvature was then drawn through the tops of these rectangles such that the area included between it and the base was the same in each interval as that of the corresponding rectangle. The base was then subdivided to represent individual years and ordinates erected to the curve so drawn. The area between the base and the curve in the interval representing each year of age then represents the number assigned to that year. And the total should agree with the total for the group.

102. Under this method, however, it was found difficult to read off the diagram with sufficient accuracy and an analytical method of redistribution has been devised. If we take the population for the various age groups and sum from the oldest group downwards we obtain a series of numbers representing the total population older than the respective ages which are the points of division between the groups. In other words the values of T_x are given for a series of values of x. If then the values of T_x can be interpolated for unit intervals we can calculate the values of L_x because we have $L_x = T_x - T_{x+1}$. Also, if the deaths are similarly treated, we have a series of values of $\lambda_x = \Sigma\theta_x$, from which, by interpolation and differencing, the successive values of θ_x may be determined. The successive values of m_x, q_x, or p_x may then be determined from the relations already given.

103. The intervals used in tabulating death returns and population statistics in the publications of the United States Census Bureau are individual years from birth to age 4 last birthday, inclusive, and five-year intervals thereafter. The returns of some countries, however, give only ten-year intervals, beginning with age 15. This grouping is adopted in order to avoid a transfer of lives from one group to another arising from a tendency to state ages at a multiple of ten years. Where the interval is ten years it is readily subdivided into five-year intervals by the finite difference formula:

$$16f(x) = 9\{f(x-t)+f(x+t)\} - \{f(x-3t)+f(x+3t)\}. \quad . \quad (11)$$

This may be easily demonstrated by expanding by Taylor's

theorem each of the functions on the right and assuming that fourth and higher differential cofficients vanish.

This formula does not apply to the last interval, where we use instead the equation

$$4\{f(x)+f(x-2t)\}=6f(x-t)+f(x-3t)+f(x+t), \quad . \quad (12)$$

which may be similarly demonstrated.

104. For the sub-division of the five-year intervals a special interpolation formula is used which ensures a continuous series. The first and second differential coefficients are determined at each point of junction by the formulas

$$12tf'(x)=8\{f(x+t)-f(x-t)\}-\{f(x+2t)-f(x-2t)\}, \quad . \quad . \quad (13)$$

$$12t^2f''(x)=16\{f(x+t)+f(x-t)\}-\{f(x+2t)+f(x-2t)\}-30f(x). \quad (14)$$

These formulas may be obtained by expansion as above, except that the fourth differential coefficient is not neglected.

For each interval a function is then found such that the values of the function itself and of its first and second differential coefficient at the beginning and end of the interval will be equal to those so determined for those points. As there are six conditions to be satisfied it follows that if a rational algebraic function is to be used it must be of the fifth degree. It may readily be demonstrated that the following function satisfies the conditions for the interval from x to $x+t$:

$$\begin{aligned} f(x+h)=f(x)&\frac{(t-h)^3(t^2+3th+6h^2)}{t^5}+f'(x)\frac{h(t-h)^3(t+3h)}{t^4} \\ &+\frac{f''(x)}{2}\frac{h^2(t-h)^3}{t^3}+f(x+t)\frac{h^3(10t^2-15th+6h^2)}{t^5} \\ &-f'(x+t)\frac{h^3(t-h)(4t-3h)}{t^4}+\frac{f''(x+t)}{2}\frac{h^3(t-h)^2}{t^3}. \quad . \quad . \quad (15) \end{aligned}$$

This may be seen by differentiating with respect to h and then putting h equal to o and t successively.

105. This equation takes a simpler form when expressed in terms of central differences as follows: Let δ denote an operation such that

$$\delta f(x) = f\left(x+\frac{t}{2}\right) - f\left(x-\frac{t}{2}\right),$$

$$\therefore \quad \delta^2 f(x) = f(x+t) - 2f(x) + f(x-t),$$

$$\delta^4 f(x) = f(x+2t) - 4f(x+t) + 6f(x) - 4f(x-t) + f(x-2t),$$

etc.

Then, from Eq. (13) we have

$$\begin{aligned} 12tf'(x) &= -f(x+2t) + 8f(x+t) - 8f(x-t) + f(x-2t), \\ &= \delta^4 f(x) - 4\delta^2 f(x) - 2\delta^2 f(x+t) + 12f(x+t) - 12f(x), \end{aligned}$$

or

$$tf'(x) = \{f(x+t) - \tfrac{1}{6}\delta^2 f(x+t)\} - \{f(x) + \tfrac{1}{3}\delta^2 f(x) - \tfrac{1}{12}\delta^4 f(x)\}. \quad (16)$$

Similarly,

$$tf'(x+t) = \{f(x+t) + \tfrac{1}{3}\delta^2 f(x+t) - \tfrac{1}{12}\delta^4 f(x+t)\} - \{f(x) - \tfrac{1}{6}\delta^2 f(x)\} \quad (17)$$

And from (14) we have similarly,

$$t^2 f''(x) = \delta^2 f(x) - \tfrac{1}{12}\delta^4 f(x), \quad . \quad . \quad . \quad . \quad . \quad (18)$$

$$t^2 f''(x+t) = \delta^2 f(x+t) - \tfrac{1}{12}\delta^4 f(x+t). \quad . \quad . \quad . \quad . \quad (19)$$

Substituting, then, these values in Eq. (15) and collecting like terms, we have, after reduction,

$$\begin{aligned} f(x+h) &= \frac{t-h}{t} f(x) - \frac{(t-h)\{t^2-(t-h)^2\}}{6t^3}\delta^2 f(x) \\ &+ \frac{h(t-h)^3(2t+5h)}{24t^5}\delta^4 f(x) + \frac{h}{t} f(x+t) - \frac{h(t^2-h^2)}{6t^3}\delta^2 f(x+t) \\ &- \frac{h^3(t-h)\{2t+5(t-h)\}}{24t^5}\delta^4 f(x+t). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20) \end{aligned}$$

106. This method cannot be applied in the above form below age 15 because $f(x-2t)$ enters into the formula and the mortality differs so much at infantile ages from that at other ages that it is not safe to assume that $f(0)$ can be determined from the same rational algebraic function as the values of $f(x)$ above age 5. Having determined, however, $f(16)$ the first

and second differential coefficients at age 10 may be determined by the equations

$$660f'(10)=216\{f(15)-f(5)\}-125\{f(16)-f(4)\}, \quad . \quad . \quad . \quad (21)$$

$$9900f''(10)=1296\{f(15)+f(5)\}-625\{f(16)+f(4)\}-1342f(10)(22)$$

These values enable us to interpolate the values of $f(x)$ for ages 11 to 14 inclusive by the use of Eq. (15). The values for ages 6 to 9 inclusive may then be filled in by determining values for $f'(5)$ and $f''(5)$ such that if 5 is put for x in Eq. (15) the values of $f(3)$ and $f(4)$ will be determined by putting h successively equal to -2 and -1.

107. Sometimes it is found preferable to interpolate by the above methods values for $\log T_x$ and $\log \lambda_x$ instead of those of T_x and λ_x, but the principle is the same. In fact, any single valued reversible function of T_x or λ_x can be used if it is found to furnish a series more appropriate for interpolation. One of the most valuable suggestions in this line is probably the use of the ratios of the values of T_x and λ_x to their values in a stationary population derived from some standard mortality table from which all minor irregularities have been removed.

The further treatment of mortality statistics of the general population will be considered under the heading of graduation, in the next chapter.

108. When dealing with the mortality experience of a life insurance company or group of such companies the problem is a different one, because in this case information is usually available as to the exact date when each life first came under observation, so that the death if it had occurred would have been included, as well as the exact date when it passed out from observation. The problem in this case is to determine the most convenient way in which the data can be analyzed and how labor can be saved without sacrificing accuracy.

It is not proposed to describe how the data can best be collected as that will depend upon various circumstances, particularly as to the mechanical sorting and tabulating devices which may be available and as to the nature of the records from which the information is to be extracted. Attention

will be confined to the principles to be adopted in classification. Aggregate mortality tables will first be dealt with, because, even where tables are constructed that are analyzed both by age and by policy duration, the differences in duration are usually neglected when the duration is in excess of some assigned limit and an aggregate table used for all longer durations.

109. The first point to which attention is directed is the analysis of the deaths, three essentially different methods having been used. The first method is known as the age year method and might otherwise be described as the exact method. Under this method the date of birth is noted and the deaths are classified precisely according to age last birthday at the time of death. In this case it is necessary to determine the number observed within each year of age and as the same life is usually observed through a series of ages the calculation can usually be most conveniently made by a continuous process, the value of E_{x+1} being determined from that of E_x by the proper modification. The particular form which the modification will take will depend on the treatment of the new entrants and withdrawals. The deaths at age x last birthday are always treated as included in E_x for the full year or as included in L_x^- for an average of half a year. In the case of new entrants and withdrawals four different methods are now available. Under the first method the exact age at entry or withdrawal may be noted and the new entrant treated as exposed for the fraction of a year of age after entry, the withdrawals being similarly treated as exposed for the fraction of a year elapsed since birthday at the time of withdrawal. Let us denote by n_x the number of new entrants at age x last birthday and by f_x the aggregate of the fractions of a year since last birthday at time of entry. Also, let w_x denote the number of withdrawals and g_x the aggregate of the fractions at withdrawal. Then we have, evidently,

$$
\begin{aligned}
E_x &= \Sigma_0^{x-1}(n_x - \theta_x - w_x) + (n_x - f_x) - (w_x - g_x), \\
&= \Sigma_0^x n_x - \Sigma_0^{x-1}\theta_x - \Sigma_0^x w_x - f_x + g_x, \\
E_{x+1} &= \Sigma_0^{x+1} n_x - \Sigma_0^x \theta_x - \Sigma_0^{x+1} w_x - f_{x+1} + g_{x+1}, \\
E_{x+1} &= E_x + (n_{x+1} - \theta_x - w_{x+1}) - (f_{x+1} - f_x) + (g_{x+1} - g_x). \quad (23)
\end{aligned}
$$

110. Under the second method of treating new entrants and withdrawals the exact fraction in each case is not calculated, but a general relation is assumed such as $f_x = fn_x$ or $g_x = gw_x$, based on an examination of part of the data taken at random. In this case we would have

$$E_{x+1} = E_x + (n_{x+1} - \theta_x - w_{x+1}) - f(n_{x+1} - n_x) + g(w_{x+1} - w_x)$$
$$= E_x - \theta_x + \{fn_x + (1-f)n_{x+1}\} - \{gw_x + (1-g)w_{x+1}\}. \quad . \quad (24)$$

Sometimes it is assumed that f and g are each equal to $\frac{1}{2}$.

In this and the preceding section those under observation at the commencement of the observations are treated as entrants at that time and those under observation at the close as withdrawals.

111. Under the third method the new entrants and withdrawals are classified according to mean age at entry or withdrawal, the mean age being calculated by deducting the calendar year of birth from the calendar year of entry or withdrawal. On the assumption that birthdays and dates of entry and withdrawal are evenly distributed over each calendar year this will give approximately correct results, the cases in which the age is overstated balancing those in which it is understated. The maximum difference between the exact age and the mean age is one year. Where the observations are closed at the end of a calendar year with a number still under observation, or started in the same way with a number already under observation, these cases must be specially treated, as the assumption of distribution of entry or exit over the calendar year does not apply. It may, however, be assumed that on an average half a year has elapsed since the last birthday. If then n_x be the new entrants at mean age x, w_x the withdrawals at the same age, σ_x those under observation at age x last birthday when the observations began, and e_x the corresponding number at the close of the observations, we have

$$E_x = \Sigma_0^x n_x - \Sigma_0^{x-1}\theta_x - \Sigma_0^x w_x + \Sigma_0^{x-1}\sigma_x + \tfrac{1}{2}\sigma_x - \Sigma_0^{x-1}e_x - \tfrac{1}{2}e_x,$$

$$E_{x+1} = E_x + \{n_{x+1} - \theta_x - w_{x+1} + \tfrac{1}{2}(\sigma_x + \sigma_{x+1}) - \tfrac{1}{2}(e_x + e_{x+1})\}. \quad (25)$$

112. Under the fourth method the age nearest birthday at entry or exit is taken instead of the mean age, the average being again correct on the assumption of uniform distribution. In this case those under observation at the opening and closing of the observations do not require special treatment, but may be grouped at age nearest birthday. If then σ_x and e_x be the numbers of such cases at age x nearest birthday, we have

$$E_x = \Sigma_0^x n_x - \Sigma_0^{x-1}\theta_x - \Sigma_0^x w_x + \Sigma_0^x \sigma_x - \Sigma_0^x e_x,$$

$$E_{x+1} = E_x + n_{x+1} - \theta_x - w_{x+1} + \sigma_{x+1} - e_{x+1}. \quad . \quad . \quad (26)$$

113. The second method of analyzing the deaths is by the policy year method under which the completed age at death is determined by adding the curtate duration (or integral part of the duration) to the age at entry determined according to whatever rule may be adopted for that purpose. Under this method the age at entry is usually taken as the age nearest birthday, but may be taken as the mean age. The age at withdrawal and of the existing at beginning or end of observations is determined by adding the duration to the age at entry. The duration may be calculated exactly in each case for the withdrawals, but it is usual to open and close the observations on policy anniversaries in chosen calendar years when the policy year method is adopted, so that there will be no fractional exposures in the case of the existing. We have, therefore,

$$E_x = \Sigma_0^x n_x - \Sigma_0^{x-1}\theta_x - \Sigma_0^x w_x + g_x + \Sigma_0^x \sigma_x - \Sigma_0^x e_x,$$

$$E_{x+1} = E_x + n_{x+1} - \theta_x - w_{x+1} + (g_{x+1} - g_x) + \sigma_{x+1} - e_{x+1}. \quad . \quad (27)$$

where w_x is the number of withdrawals at completed age x determined as for the deaths, and g_x is the aggregate of the fractional durations. Approximate methods of calculating the values of g_x are sometimes used.

The duration of the withdrawals may also be taken as an exact number of years in each case, the mean duration or the nearest duration being used. When this is done fractional durations disappear and we see that in this case Eq. (26) holds.

114. The third method of analyzing the deaths is by calendar years, the completed age at death being the age nearest birthday at the beginning of the calendar year of death. Under this method the ages at entry and withdrawal may be determined by any of the methods outlined for the policy year method and the equations of relation will be the same as for that method. In many cases where the calendar year method is applied the age at the beginning of the calendar year is not taken as the age nearest birthday, but is calculated by adding the curtate duration plus half a year to the age nearest birthday at entry or by adding the curtate duration to the age next birthday at entry. In the former case the ages for which the rates of mortality are determined will not be integral, but will be of the form $x+\frac{1}{2}$, where x is integral.

115. To summarize, let x be the exact age at entry, $\underline{x}|$ the age last birthday, so that the age next birthday is $\underline{x}|+1$, (x) the age nearest birthday, $|x|$ the mean age, and $[x]$ the the assumed age at entry. Also let t be the exact duration, $\underline{t}|$ the curtate duration, (t) the nearest integral duration, and $|t|$ the mean duration. Then, under the age year methods the completed age at death is taken as $\underline{x+t}|$, under the policy year methods it is taken as $[x]+\underline{t}|$, and under the calendar year methods it is taken as either $[x]+|t|-\frac{1}{2}$, or $|x+t|-\frac{1}{2}$ or $|x-\frac{1}{2}|+|t|$. In the last expression $|x-\frac{1}{2}|$ is used to designate the mean age six months before entry. In connection with the age at entry we may have $[x]$ taken as x, $\underline{x}|+f$, (x) or $|x|$, and in connection with the withdrawals the age at withdrawal may be taken as $x+t$, $(x+t)$, $|x+t|$, $\underline{x+t}|+g$, $[x]+t$, $[x]+(t)$, $[x]+|t|$ or $[x]+\underline{t}|+g$.

116. The census methods may also be conveniently applied to a life insurance company's experience, a classification of the lives insured being made at the close of each calendar year, giving the population in each year of age at that time. The completed age may be taken as $\underline{x+t}|$ or $[x]+\underline{t}|$, usually the latter. If, then, the deaths of a given calendar year be analyzed by completed age at death, the average population during that year for any age year can be determined by taking

the mean of the populations at the beginning and end of the year. The exposed to risk will then be determined by adding to this mean population one-half of the deaths. Each successive calendar year is treated in the same way and the experience gradually accumulated, the experience of as many years as desired being combined.

By a modification of this method the deaths are observed from the middle of one calendar year to the middle of the next, or from the policy anniversaries in one calendar year to those in the next, and the population at the end of the first year is taken to represent the corresponding average population.

117. Where an analyzed mortality table is to be constructed the policy year method is usually adopted and the experience during each year of duration is treated separately so far as is considered necessary. This amounts to the same thing as treating each age at entry separately for as many years of duration as is decided upon. This latter method is the more convenient for descriptive purposes, because new entrants are thereby eliminated after the start and we need consider only deaths, withdrawals, existing at the beginning of observations and existing at end. Assuming that experience before the policy anniversary in the first calendar year or after it in the last calendar year of the observations is neglected and that nearest durations are taken for withdrawals, we have

$$E_{[x]+t} = n_{[x]} + \Sigma_1^t \sigma_{[x]+t} - \Sigma_0^t w_{[x]+(t)} - \Sigma_1^t e_{[x]+t} - \Sigma_0^{t-1} \theta_{[x]+\underline{t}|},$$

$$E_{[x]+t+1} = E_{[x]+t} + \sigma_{[x]+t+1} - w_{[x]+(t+1)} - e_{[x]+t+1} - \theta_{[x]+\underline{t}|}. \qquad (28)$$

$$q_{[x]+t} = \theta_{[x]+\underline{t}|} / E_{[x]+t} \qquad (29)$$

118. Having obtained the values of $q_{[x]+t}$, those of $l_{[x]+t}$, and $d_{[x]+t}$ are obtained as follows: The ultimate mortality table representing the mortality after the effect of selection is assumed to have disappeared is first constructed in the way indicated in Chapter II by taking an initial value of l_x and multiplying successively by the successive values of p_x. The values of $l_{[x]+t}$ are then constructed in reverse order,

beginning with the value of t at which the analyzed mortality merges into the ultimate. The working formula is

$$l_{[x]+t} = l_{[x]+t+1} \div p_{[x]+t}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

The values of $d_{[x]+t}$ are then derived by differencing from the relation

$$d_{[x]+t} = l_{[x]+t} - l_{[x]+t+1}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

119. Diagram No. 1 is shown to illustrate the relation between the rates of mortality in an aggregate table and in a select or analyzed table based on the same data. In this diagram, in order to avoid the necessity of a change of scale, the ordinates are made proportionate to log $(1+100q_x)$, instead of to q_x. A scale is given on the margin, however, showing the values of q_x corresponding to different lengths of ordinates. The tables selected to illustrate this point are the O^M and $O^{[M]}$ tables. It will be noted that at the young ages the rates of mortality in the aggregate table approach those shown in the analyzed table for the first year, whereas at the older ages they approach and become indistinguishable from those in the ultimate part of the analyzed table.

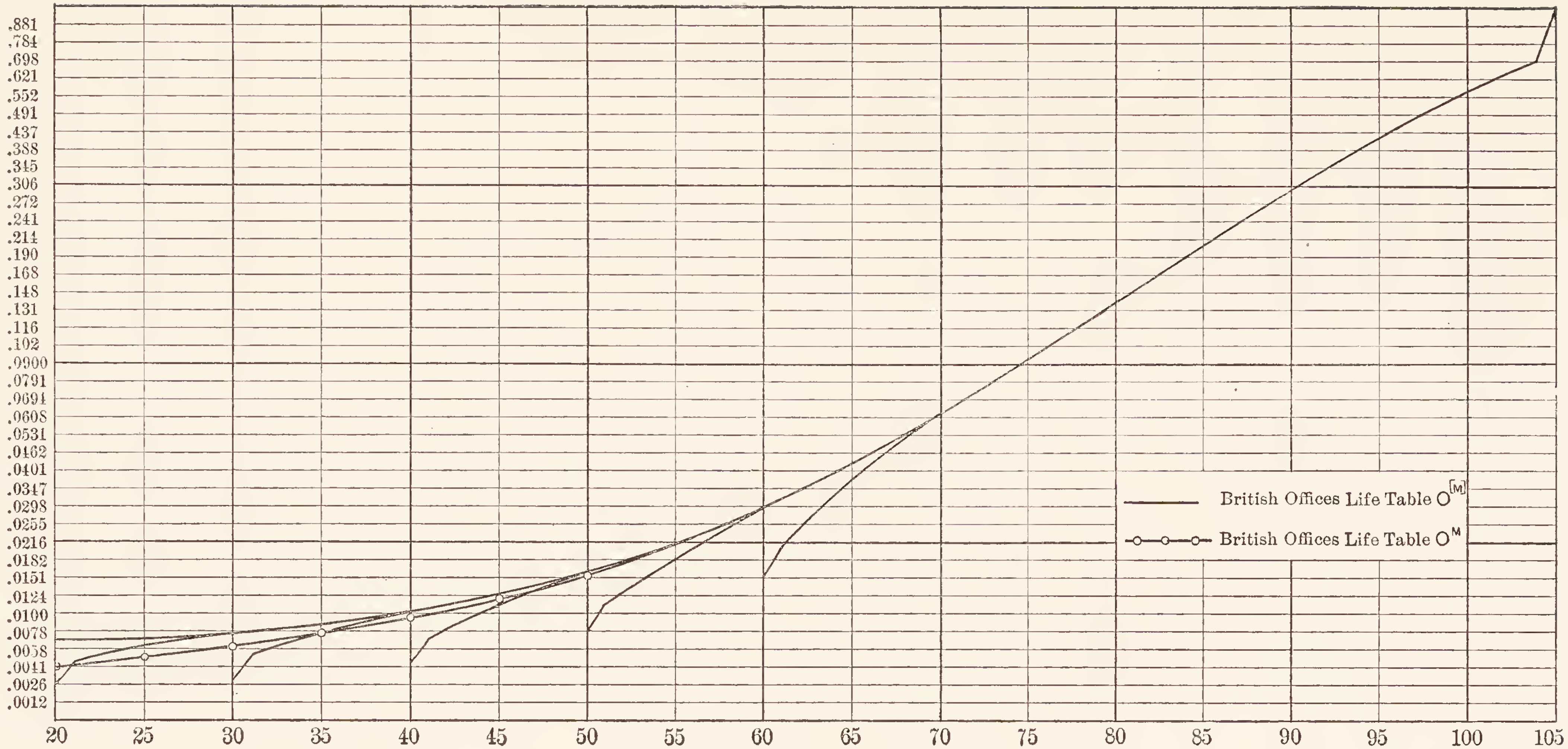

DIAGRAM NO. 1.—COMPARISON OF AGGREGATE AND ANALYZED RATES OF MORTALITY.

The Abscissas represent Ages. The Ordinates are proportional to values of $\log(1+100\,q_x)$, where q_x is the probability of dying at age x, and the numbers on the left margin show values of q_x.

CHAPTER VII

GRADUATION OF MORTALITY TABLES

120. When the probability of dying within a year is q_x and n lives are exposed to the risk of death, the expected number of deaths is nq_x. This means that in a large number of such instances, where n lives are observed in each instance, the average number of deaths occurring will be nq_x. It does not mean, however, that in every instance exactly nq_x deaths will occur. In fact, unless the values of n and q_x happen to be so related that nq_x is an integer, the exact relation cannot hold. Any number of deaths from none up to n is theoretically possible. The probability of exactly r deaths is shown by the principles of the theory of probability to be $\frac{|\underline{n}}{|\underline{r}\,|\underline{n-r}} q_x^r p_x^{n-r}$ and in that case the deviation of the actual number of deaths from the expected will be $r-nq_x$. The mean value of the square of this deviation may be shown to be np_xq_x. Where n is large this deviation is equally likely to be positive or negative and there is approximately an even chance that it will exceed $\frac{2}{3}\sqrt{np_xq_x}$ in absolute magnitude. If, then, a value $q'_x=\frac{r}{n}$ be determined by dividing the observed number of deaths by the number exposed, there is an even chance that this value will differ from the true probability q_x by at least as much as $\frac{2}{3}\sqrt{\frac{p_xq_x}{n}}$, and the deviation is equally likely to be positive or negative. We have no means, however, of determining from the observations themselves at age x whether q'_x is greater or less than q_x and in the absence of further

information the hypothesis is adopted that $q_x = q'_x$, because that is the value of q_x which makes the probability $\frac{|\underline{n}}{|\underline{r}\,|\underline{n-r}} p_x^{n-r} q_x^r$ of the observed facts a maximum.

121. If, now, we have made similar observations at a series of consecutive ages for which the true probabilities of death are q_x, q_{x+1}, etc., the principle of continuity would lead us to expect that the successive values of q_x for the different ages would form a smooth series. Where the values of q'_x are calculated each will differ from the corresponding value of q_x by a quantity which may be large or small, positive or negative, a positive deviation being as likely to be followed by a negative one as by a positive. The theory, however, indicates that the probable values of these deviations in the value of q_x decrease as n increases and that positive and negative values tend to counterbalance one another. It follows from the above that we must expect the values of q'_x to form a somewhat irregular series, the amount of the irregularity depending on the numbers under observation. The problem of graduation is to remove the irregularities from this series and to approximate as closely as possible to the true values of q_x.

122. The methods which have been adopted for this purpose come under four general classes. Under the first method a diagram is made to represent graphically the observed facts and a continuous curve is then drawn as a basis for the graduated series. Under the second method the graduated series is formed by interpolation on the basis of values determined for fixed intervals, these values being so determined as to give an interpolated series fitting as closely as possible to the observed facts. Under the third method the individual terms of the graduated series are each determined by a summation of adjacent terms of the original series, a correction being introduced to allow for the curvature. Under the fourth method a mathematical formula containing arbitrary constants is used to express the series and the constants are determined so as to agree in certain respects with the observed facts.

123. After a graduated series has been constructed it is usually tested with respect to the two points of smoothness and closeness to the observed facts. With respect to smoothness the fact that a series is determined by a mathematical formula is usually taken as sufficient, but when it is not so determined the criterion usually adopted is the smallness of the third differences in the graduated series. This smallness is sometimes tested by inspection of the differences after they have been taken out, but in comparing two different graduations of the same series, if it is desired to have a numerical measure of their departure from absolute smoothness, the sum of the squares of the third differences or the sum of the absolute values irrespective of sign of such differences may be taken as such measure.

124. With respect to closeness to the observed facts the requirements usually made are (1) that the total number of expected deaths and their first and second moments about any assigned age shall be the same as for the actual deaths, and (2) that the departures in individual groups shall not, on the average, materially exceed in magnitude those expected in accordance with the theory of probability. This comparison is usually made by recording the difference between the actual and expected deaths at each age. A continuous summation of these deviations is then made with due regard to sign. The smallness of the numbers in this column of accumulated deviations, the frequency of changes of sign and the extent to which positive and negative terms balance one another indicate the extent to which the first requirement is complied with. The sum of the deviations without regard to sign tests directly the second requirement if the average deviation which is approximately equal to $\frac{4}{5}\sqrt{np_xq_x}$ be calculated for each age for the purpose of comparison. The comparison may also be based on the sum of the squares of the departures or on the sum of those squares each divided by its mean value np_xq_x. The test in this last form is supported logically by the fact that if the number observed is large the quantity so arrived at is proportional to the logarithm of the ratio between the

probability of the observed facts and that of the expected according to the graduated table.

125. The graphic method of graduating mortality tables arises naturally from the graphic method of representing them. Under this method the values of q_x or of m_x are represented by points in a diagram. For convenience in plotting the diagram accurately ruled section paper is ordinarily used, the years of age being represented by equal intervals along the base line and the rate being represented on a suitable scale by the distance of the point from the base. When the points corresponding to the successive ages are plotted and joined by straight lines it is found in an ungraduated table that the result is a zigzag line full of minor irregularities, but showing indications, the strength of which depends on the volume of the observations, of an underlying regular law. The graduation of the table is effected by drawing among these points, but not necessarily through any of them, a regular curve to represent this law. Preliminary groupings not covering equal intervals, but so arranged as to produce the greatest attainable regularity are made in order to bring out this law. After the curve is drawn, the values of the ordinates are read off and the results corrected to remove any irregularities due to errors in reading. A comparison is then made between the expected and actual deaths on the lines indicated above and, if a relatively large and persistent deviation in either direction is accumulated in any section of the table, the curve is amended to reduce or eliminate it.

126. In applying this method the difficulty is found that, if the scale of the diagram is sufficiently large to permit of accurate reading in one part of the curve, it will be too large in another. This difficulty is met by plotting not the actual value of q_x but some more slowly varying quantity from which it can be determined. One method is to take as a basis some mortality table already constructed from a mathematical formula such as Makeham's and to plot either the ratio of the observed q_x to the rate at the same age in the table, or the difference of the rates. Another method is to use $\log(1+100q_x)$

or $\log(1+10q_x)$ instead of q_x. Specially ruled section paper has been prepared with the vertical spacing so arranged that if the values of q_x are plotted according to the ruling the actual distances from the base will be $\log(1+100q_x)$, so that, if this paper is used, the whole diagram is reduced to practicable dimensions without altering the scale where q_x is small, and after the curve is drawn the values of q_x may be read off directly.

127. In the interpolation method the problem may be divided into three parts. The first is to determine which function of the mortality to interpolate, the second is to determine the values for the selected ages of that function and the third part is to interpolate the intermediate values. Various functions have been used as a basis for interpolation and the method of determining the values at the selected ages will depend on the function selected. Where l_x or $\log l_x$ is used we must have an ungraduated mortality table as a basis and we obtain our points for interpolation by simply taking the values at the selected intervals from this ungraduated table. Where, however, q_x or some function of q_x, such as $\log q_x$, m_x, or $\log m_x$, is selected for interpolation we must group the data so as to determine with as much precision as possible the value of the function for the selected ages. For this purpose the exposed to risk, or the population, and the deaths are usually combined into age groups. In the case of the earlier English Life Tables the simple but somewhat inaccurate assumption was made that the total deaths in an age group of five or ten years divided by the total population corresponding to them would give the force of mortality at the central age of the group as explained in Chapter VI.

128. Where more accurate values are desired the redistributed values of the deaths and exposures or population for some one year of age in each group may be calculated by methods similar to those described for census statistics in the preceding chapter. For this purpose quinquennial age groups are usually used, and in the case of general population statistics these groupings are given to us ready made or if not can only be formed by subdividing decennial groups. In the case, how-

ever, of the experience of insurance companies the groups are formed by combining the figures for individual ages and we have freedom of choice as to the limits of the groups.

129. If the central year of each quinquennial group is taken, it is usual to assume that the population and the deaths for the successive years may be expressed by a rational algebraic function of the third degree, so that we have, for each function, an equation of the form

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\frac{h^3}{6}f'''(x).$$

Summing this from $h=-2$ to $h=+2$ inclusive we get for the total number included in the group of which x is the central age, $5f(x)+5f''(x)$. This may be expressed by the equation $w_x=5f(x)+5f''(x)$, if we use w_x to denote the total of the group.

Similarly, summing from $h=-7$ to $h=+7$ inclusive we have

$$w_{x-5}+w_x+w_{x+5}=15f(x)+140f''(x).$$

Eliminating $f''(x)$ between these two equations, we have

$$\begin{aligned}125f(x)&=27w_x-(w_{x-5}+w_{x+5})=25w_x-(w_{x-5}-2w_x+w_{x+5})\\&=25w_x-\delta^2 w_x,\end{aligned}$$

or

$$5f(x)=w_x-\tfrac{1}{25}\delta^2 w_x. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

It thus appears that the adjustment to be made to the total of the groups to obtain the number for the central year is the same in form for the deaths, the population, or the exposed to risk.

130. Where the first year of each group is taken, it is necessary to use four age groups in order to determine the value of $f(x)$. We have

$$\begin{aligned}w_{x-8}&=\Sigma_{h=6}^{h=10}\left\{f(x)-hf'(x)+\frac{h^2}{2}f''(x)-\frac{h^3}{6}f'''(x)\right\},\\&=5f(x)-40f'(x)+165f''(x)-\frac{2800}{6}f'''(x),\end{aligned}$$

$$w_{x-3}=\Sigma_{h=1}^{h=5}\left\{f(x)-hf'(x)+\frac{h^2}{2}f''(x)-\frac{h^3}{6}f'''(x)\right\},$$

$$=5f(x)-15f'(x)+\frac{55}{2}f''(x)-\frac{225}{6}f'''(x),$$

$$w_{x+2}=\Sigma_{h=0}^{h=4}\left\{f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\frac{h^3}{6}f'''(x)\right\},$$

$$=5f(x)+10f'(x)+15f''(x)+\frac{100}{6}f'''(x),$$

$$w_{x+7}=\Sigma_{h=5}^{h=9}\left\{f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\frac{h^3}{6}f'''(x)\right\},$$

$$=5f(x)+35f'(x)+\frac{255}{2}f''(x)+\frac{1925}{6}f'''(x),$$

$$2w_{x-3}+3w_{x+2}=25f(x)+100f''(x)-25f'''(x),$$

$$\delta^2w_{x-3}=125f''(x)-375f'''(x),$$

$$\delta^2w_{x+2}=125f''(x)+250f'''(x),$$

$$9\delta^2w_{x-3}+11\delta^2w_{x+2}=2500f''(x)-625f'''(x),$$

$$\therefore\quad 625f(x)=25(2w_{x-3}+3w_{x+2})-(9\delta^2w_{x-3}+11\delta^2w_{x+2}),$$

$$f(x)=\frac{2w_{x-3}+3w_{x+2}}{25}-\frac{9\delta^2w_{x-3}+11\delta^2w_{x+2}}{625}, \quad . \quad (2)$$

131. On account of the lack of symmetry of the above formula, it may be considered desirable to select a year of age, one-half of which will be in one age group and the other half in the next group. In this case and in cases where the force of mortality is to be determined it is more convenient to work with a function $\phi(x)$ representing the number per unit of age at exact age x, so that

$$f(x)=\int_0^1\phi(x+h)dh,$$

and

$$w_{x+2}=\int_0^5\phi(x+h)dh,$$

or

$$w_{x-\frac{1}{2}}=\int_{-2\frac{1}{2}}^{+2\frac{1}{2}}\phi(x+h)dh.$$

Expanding $\phi(x+h)$ and integrating, we have

$$w_{x-8}=\int_{-10}^{-5}\phi(x+h)dh=5\phi(x)-\frac{75}{2}\phi'(x)+\frac{875}{6}\phi''(x)-\frac{9375}{24}\phi'''(x)$$

$$w_{x-3}=\int_{-5}^{0}\phi(x+h)dh=5\phi(x)-\frac{25}{2}\phi'(x)+\frac{125}{6}\phi''(x)-\frac{625}{24}\phi'''(x)$$

$$w_{x+2}=\int_{0}^{5}\phi(x+h)dh=5\phi(x)+\frac{25}{2}\phi'(x)+\frac{125}{6}\phi''(x)+\frac{625}{24}\phi'''(x)$$

$$w_{x+7}=\int_{5}^{10}\phi(x+h)dh=5\phi(x)+\frac{75}{2}\phi'(x)+\frac{875}{6}\phi''(x)+\frac{9375}{24}\phi'''(x)$$

$$w_{x-3}+w_{x+2}=10\phi(x)+\frac{125}{3}\phi''(x),$$

$$w_{x-8}+w_{x+7}=10\phi(x)+\frac{875}{3}\phi''(x).$$

But

$$f(x-\tfrac{1}{2})=\int_{-\frac{1}{2}}^{+\frac{1}{2}}\phi(x+h)dh=\phi(x)+\frac{1}{24}\phi''(x),$$

and

$$\delta^2 w_{x-3}+\delta^2 w_{x+2}=(w_{x-8}-2w_{x-3}+w_{x-2})+(w_{x-3}-2w_{x+2}+w_{x+7}),$$
$$=(w_{x-8}+w_{x+7})-(w_{x-3}+w_{x+2})=250\phi''(x)$$

$$\therefore\quad 10f(x-\tfrac{1}{2})=(w_{x-3}+w_{x+2})-\frac{495}{12}\phi''(x),$$
$$=(w_{x-3}+w_{x+2})-.165(\delta^2 w_{x-3}+\delta^2 w_{x+2}),$$
$$=(w_{x-3}-.165\,\delta^2 w_{x-3})+(w_{x+2}-.165\,\delta^2 w_{x+2}). \quad (3)$$

This covers the case where half of the year of age is in each group.

132. Where the force of mortality is required at the age corresponding to the point of division we have

$$60\phi(x)=7(w_{x-3}+w_{x+2})-(w_{x-8}+w_{x+7}),$$

or

$$10\phi(x)=(w_{x-3}-\tfrac{1}{6}\delta^2 w_{x-3})+(w_{x+2}-\tfrac{1}{6}\delta^2 w_{x+2}), \quad . \quad . \quad (4)$$

which is the formula to be used for the deaths and population.

133. For the age corresponding to the center of the group we have

$$w_{x-\frac{1}{2}} = \int_{-2\frac{1}{2}}^{+2\frac{1}{2}} \phi(x+h)dh = 5\phi(x) + \frac{125}{24}\phi''(x),$$

$$w_{x-5\frac{1}{2}} + w_{x-\frac{1}{2}} + w_{x+4\frac{1}{2}} = \int_{-7\frac{1}{2}}^{+7\frac{1}{2}} \phi(x+h)dh = 15\phi(x) + \frac{3375}{24}\phi''(x),$$

whence

$$120\phi(x) = 26w_{x-\frac{1}{2}} - (w_{x-5\frac{1}{2}} + w_{x+4\frac{1}{2}}),$$

$$= 24w_{x-\frac{1}{2}} - \delta^2 w_{x-\frac{1}{2}},$$

or

$$5\phi(x) = w_{x-\frac{1}{2}} - \frac{1}{24}\delta^2 w_{x-\frac{1}{2}}. \qquad (5)$$

134. These adjustments enable us to calculate the values of q_x, m_x, or μ_x for values of x separated by quinquennial ages, and the remainder of the problem consists in interpolating intermediate values. For this purpose a formula of osculatory interpolation will be found most satisfactory. Three such formulas have been proposed. The first, which is the basis of the other two, is the one described in Chapter VI in connection with the redistribution of population and deaths, the simplest form for practical application being that given in Eq. (20) of that chapter.

$$\begin{aligned} f(x+h) &= \frac{t-h}{t}f(x) - \frac{(t-h)\{t^2-(t-h)^2\}}{6t^3}\delta^2 f(x) \\ &+ \frac{h(t-h)^3(2t+5h)}{24t^5}\delta^4 f(x) + \frac{h}{t}f(x+t) - \frac{h(t^2-h^2)}{6t^3}\delta^2 f(x+t) \\ &+ \frac{(t-h)h^3(7t-5h)}{24t^5}\delta^4 f(x+t). \qquad (6) \end{aligned}$$

The other two formulas are simplifications of this by omitting some of the conditions, the first differential coefficient only being determined for the points of junction.

135. In Karup's form, which is the simpler, this first differential coefficient is determined on the assumption of a curve of the second degree for $f(x)$, the value so derived being $2tf'(x) = f(x+t) - f(x-t)$. A curve of the third degree is then

determined so as to have the required values and first differential coefficients at the points of junction. The equation of this curve is found to be

$$f(x+h)=\frac{t-h}{t}f(x)-\frac{h(t-h)^2}{2t^3}\delta^2 f(x)$$
$$+\frac{h}{t}f(x+t)-\frac{h^2(t-h)}{2t^3}\delta^2 f(x+t). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

136. Greater accuracy without material increase of work can be obtained by determining the first differential coefficient on the assumption of a fourth difference curve by Eq. (13) of Chapter VI, namely,

$$12tf'(x)=8\{f(x+t)-f(x-t)\}-\{f(x+2t)-f(x-2t)\}.$$

This condition is then found to be satisfied by the equation

$$f(x+h)=\frac{t-h}{t}\left\{f(x)-\frac{h}{6t}\left[\delta^2 f(x)+\frac{t-h}{t}(\delta^2 f(x)-\tfrac{1}{2}\delta^4 f(x))\right]\right\}$$
$$+\frac{h}{t}\left\{f(x+t)-\frac{t-h}{6t}\left[\delta^2 f(x+t)+\frac{h}{t}(\delta^2 f(x+t)-\tfrac{1}{2}\delta^4 f(x+t))\right]\right\}. \quad (8)$$

137. When making this interpolation it is usually necessary to assume the age at which q_x becomes equal to unity and the age chosen should be consistent with the data at the old ages. When it has been chosen it will be advisable to arrange the division into groups, if possible, in such a way that the ages for which the function is determined will form with the limiting age a regular series of differences. For example, if it is assumed that $q_{102}=1$ or $m_{102}=2$, then we should so arrange the groupings as to determine the rates for ages ending in 2 and 7, starting the groups at those ages if it is intended to use the first year of each group and starting the groups at ages ending in 0 or 5 if the central age of each group is to be used.

138. If we have interpolated a series of values of μ_x for each age, then we can pass to values of p_x by the approximate formula

$$\operatorname{colog}_e p_x = \int_x^{x+1} \mu_x dx$$

$$= \mu_x + \frac{1}{2}\frac{d\mu_x}{dx} + \frac{1}{6}\frac{d^2\mu_x}{dx^2} + \frac{1}{24}\frac{d^3\mu_x}{dx^3} + \text{etc.},$$

$$= \frac{13(\mu_x + \mu_{x+1}) - (\mu_{x-1} + \mu_{x+2})}{24}. \quad . \quad . \quad . \quad (9)$$

139. The summation methods of graduation can be applied only when we have constructed a complete table of the ungraduated values of the function to be graduated. In investigating the effect of a graduation formula we may consider the ungraduated value of a function as consisting of two parts, one the true value V_x of the function which would result from an indefinitely extended experience and the other the error or deviation E_x of the observed value from the true value. The fundamental assumption which is the basis of all graduation is that the values of V_x form a regular or smooth series and that the values of E_x form an irregular series, the fluctuations in the value of any one term being independent of those of the neighboring terms. It is also assumed that the mean value of each error, E_x, is zero when the sign of the error is taken into account.

140. Let us first assume that over a short range of values of x, the values of V_x may be considered as forming an arithmetic series. Then it is evident that if we take the arithmetic mean of an odd number of terms of the series of values of V_x it will be exactly equal to the middle term. Consequently, if we take a similar average of the ungraduated values of the function it will differ from the true value V_x corresponding to the middle term by the average of the values of E_x, and it follows from principles of averages that the mean absolute value of this resultant deviation will be much smaller than that of the individual ungraduated values. In this case, therefore, a simple average of a number of terms will constitute a graduation of the series. It is evident that the same process may be repeated without disturbing the value of V_x, but the effect on the values of E_x diminishes with the successive summation

because the neighboring values are no longer independent, each original deviation now affecting, although to a smaller extent, a number of successive values. For example, suppose we take the average of five successive terms, then the error in the resulting average will be $\frac{1}{5}(E_{x-2}+E_{x-1}+E_x+E_{x+1}+E_{x+2})$. If we assume that the mean value of the square of each of these primary errors is μ_2, then it is evident that the mean value of the square of the error of the averages will be $\frac{1}{5}\mu_2$. If we repeat the process a second time, the expression for the error will be

$$\tfrac{1}{25}(E_{x-4}+2E_{x-3}+3E_{x-2}+4E_{x-1}+5E_x+4E_{x+1}+3E_{x+2} \\ +2E_{x+3}+E_{x+4}),$$

the mean value of the square of which is

$$\frac{85}{25^2}\mu_2=\frac{17}{125}\mu_2=\frac{17}{25}\cdot\frac{1}{5}\mu_2.$$

If it is repeated a third time the error is

$$\tfrac{1}{125}\{E_{x-6}+3E_{x-5}+6E_{x-4}+10E_{x-3}+15E_{x-2}+18E_{x-1}+19E_x+18E_{x+1} \\ +15E_{x+2}+10E_{x+3}+6E_{x+4}+3E_{x+5}+E_{x+6}\},$$

the mean value of the square of which is

$$\frac{1751}{125^2}\mu_2=\frac{103}{125}\cdot\frac{17}{25}\cdot\frac{1}{5}\mu_2.$$

141. Let us, however, investigate the effect on the third differences of the series. On the assumption made, that the values of V_x form an arithmetic series, the third differences will arise entirely from the errors, the expression for the third differences in the ungraduated series being

$$E_{x+3}-3E_{x+2}+3E_{x+1}-E_x,$$

the mean value of the square of which is $20\mu_2$. When the first average is taken the third difference becomes

$$\tfrac{1}{5}(E_{x+5}-2E_{x+4}+E_{x+3}-E_x+2E_{x-1}-E_{x-2})$$

the mean value of the square of which is

$$\frac{12}{25}\mu_2=\frac{3}{125}\cdot 20\mu_2.$$

For the second average the third difference is

$$\tfrac{1}{25}\{E_{x+7}-E_{x+6}-2E_{x+2}+2E_{x+1}+E_{x-3}-E_{x-4}\},$$

the mean value of the square of which is

$$\frac{12}{625}\mu_2=\frac{1}{25}\cdot\frac{3}{125}\cdot 20\mu_2.$$

For the third average the third difference is

$$\tfrac{1}{125}\{E_{x+9}-3E_{x+4}+3E_{x-1}-E_{x-6}\},$$

the mean value of the square of which is

$$\frac{20}{125^2}\mu_2=\frac{1}{15}\cdot\frac{1}{25}\cdot\frac{3}{125}\cdot 20\mu_2.$$

It will be noted that for successive summations the effect on the smoothness of the series of errors does not diminish as rapidly as the effect on the absolute values.

142. The summations do not necessarily all extend over the same number of terms nor is the number of terms in each summation necessarily odd because, while an average over an even number of terms of a series does not give a term of the series but a term midway between two of them, a second average over the same or a different even number of terms will bring us back to the original series. An even number of summations over an even number of terms each may therefore be introduced and the resulting averages will correspond to terms of the original series. For example, suppose, instead of taking the average in fives three times we take the average in fours, fives and sixes. Then the expression for the resulting error will be

$$\tfrac{1}{120}(E_{x-6}+3E_{x-5}+6E_{x-4}+10E_{x-3}+14E_{x-2}+17E_{x-1}+18E_x+17E_{x+1}$$
$$+14E_{x+2}+10E_{x+3}+6E_{x+4}+3E_{x+5}+E_{x+6}),$$

the mean value of the square of which is $\frac{1586}{120^2}\mu_2$, or a little less than for the third average in fives. The expression for the third difference is

$$\tfrac{1}{120}(E_{x+9}-E_{x+5}-E_{x+4}-E_{x+3}+E_x+E_{x-1}+E_{x-2}-E_{x-6});$$

the mean value of the square of which is $\frac{8}{120^2}\mu_2$. We thus see that by making the periods unequal a slight increase in weight is obtained on the individual term and a great increase in weight in the third difference. In other words, with unequal summations a much smoother series is obtained and it is at the same time a little more accurate.

143. Thus far we have assumed that the series of true values is an arithmetic series, the general term of which may be expressed by a linear function. It is necessary, however, to take into account the second and higher differential coefficients. The summation graduation formulas ordinarily used contain a correction so that a series the general term of which is of the third degree will be reproduced. Where the function is of the third degree, we have, generally,

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\frac{h^3}{6}f'''(x).$$

If, then, we add together n terms of this series, we have

$$f(x)+f(x+1)+\ldots f(x+n-1)$$
$$=nf(x)+\frac{n(n-1)}{2}f'(x)+\frac{n(n-1)(2n-1)}{12}f''(x)+\frac{n^2(n-1)^2}{24}f'''(x).$$

But

$$nf\left(x+\frac{n-1}{2}\right)=nf(x)+\frac{n(n-1)}{2}f'(x)$$
$$+\frac{n(n-1)^2}{8}f''(x)+\frac{n(n-1)^3}{48}f'''(x),$$

$$\frac{n(n^2-1)}{24}f''\left(x+\frac{n-1}{2}\right)=\frac{n(n^2-1)}{24}f''(x)+\frac{n(n^2-1)(n-1)}{48}f'''(x).$$

$$\therefore\quad f(x)+f(x+1)+\ldots f(x+n-1)=nf\left(x+\frac{n-1}{2}\right)$$
$$+\frac{n(n^2-1)}{24}f''\left(x+\frac{n-1}{2}\right).$$

If, then, we denote

$$f\left(x-\frac{n-1}{2}\right)+f\left(x-\frac{n-3}{2}\right)+\ldots f\left(x+\frac{n-1}{2}\right),$$

by $[n]f(x)$, we have, generally,

$$[n]f(x)=nf(x)+\frac{n(n^2-1)}{24}f''(x),$$

or

$$\frac{[n]f(x)}{n}=f(x)+\frac{n^2-1}{24}f''(x).$$

144. The operation of taking the average of n successive terms has therefore introduced an error $\frac{n^2-1}{24}f''(x)$ into the value of $f(x)$. If, then, we repeat the process with the same or a different number of terms, an additional error of the same form is introduced. We also see that, since according to our assumptions $f^{\mathrm{IV}}(x)$ vanishes for all values of x and consequently $f''(x)$ is linear in form, the error introduced by the first average is carried forward unchanged. If, then, three successive averages cover p, q, and r terms, we have,

$$\frac{[p][q][r]f(x)}{pqr}=f(x)+\left\{\frac{p^2-1}{24}+\frac{q^2-1}{24}+\frac{r^2-1}{24}\right\}f''(x),$$

$$=f(x)+\frac{p^2+q^2+r^2-3}{24}f''(x).$$

Where the number of terms in each average is the same and each equal to n this may be expressed as follows:

$$\frac{[n]^3f(x)}{n^3}=f(x)+\frac{n^2-1}{8}f''(x).$$

145. A correction must accordingly be introduced to compensate for this error. If, then, we use $\gamma_n f(x)$ as a short expression for $f(x+n)+f(x-n)$ we have

$$\gamma_n f(x)=2f(x)+n^2f''(x),$$

or

$$\gamma_n f(x)-2f(x)=n^2f''(x).$$

we have, therefore,

$$(a\gamma_1+b\gamma_2+c\gamma_3)f(x)=2(a+b+c)f(x)+(a+4b+9c)f''(x),$$

or

$$(1+2a+2b+2c)f(x)-(a\gamma_1+b\gamma_2+c\gamma_3)f(x)$$
$$=f(x)-(a+4b+9c)f''(x).$$

If, therefore, we substitute,

$$(1+2a+2b+2c)f(x)-(a\gamma_1+b\gamma_2+c\gamma_3)f(x) \text{ for } f(x),$$

before averaging, we have

$$\frac{[p][q][r]}{pqr}\{(1+2a+2b+2c)-(a\gamma_1+b\gamma_2+c\gamma_3)\}f(x)$$

$$=\frac{[p][q][r]}{pqr}\{f(x)-(a+4b+9c)f''(x)\}$$

$$=f(x)+\left\{\frac{p^2+q^2+r^2-3}{24}-(a+4b+9c)\right\}f''(x).$$

If then a, b, and c are so determined that

$$a+4b+9c=\frac{p^2+q^2+r^2-3}{24},$$

a series of the third degree will be exactly reproduced. We have here three unknowns and only one condition, so that the equation is an indeterminate one and the various formulas result from the use of different values of p, q, and r and of a, b, and c, the condition in some cases not being exactly satisfied.

146. The first summation formula correct to third differences was that devised by Woolhouse. It was not originally constructed as a summation formula, but it was afterwards found to take that form. In this formula $p=q=r=5$, so that

$$\frac{p^2+q^2+r^2-3}{24}=3.$$

Also $a=3$, and $b=c=0$, so that the formula takes the form $\frac{[5]^3}{5^3}\{7-3\gamma_1\}f(x)$. Woolhouse's formula when expanded as a function of the terms of the series becomes

$$\tfrac{1}{125}(25+24\gamma_1+21\gamma_2+7\gamma_3+3\gamma_4-2\gamma_6-3\gamma_7)f(x).$$

147. J. A. Higham, who first showed that Woolhouse's formula might be applied by the summation method, suggested an alternative compensating adjustment which was equiv-

alent to putting $a=-1$, $b=1$ and $c=0$, thus still keeping $a+4b+9c=3$. His formula therefore took the form

$$\frac{[5]^3}{5^3}(1+\gamma_1-\gamma_2)f(x)=\frac{[5]^3}{5^3}([3]-\gamma_2)f(x),$$

which expands into

$$\tfrac{1}{125}(25+24\gamma_1+18\gamma_2+10\gamma_3+3\gamma_4-2\gamma_6-2\gamma_7-\gamma_8)f(x).$$

148. Karup's graduation formula uses summations in fives with the compensating adjustment $a=-\frac{3}{5}$, $b=0$, $c=\frac{2}{5}$, so that the formula becomes,

$$\frac{[5]^3}{5^3}(\tfrac{3}{5}+\tfrac{3}{5}\gamma_1-\tfrac{2}{5}\gamma_3)f(x)=\frac{[5]^3}{5^4}(3[3]-2\gamma_3)f(x).$$

149. G. F. Hardy used the same compensating adjustment as Higham, but used successive summations in fours, fives, and sixes, so that for his formula $\dfrac{p^2+q^2+r^2-3}{24}=3\tfrac{1}{12}$, and it is not fully compensated, the remaining second difference error being $\frac{1}{12}f''(x)$. This is, however, small and is approximately counterbalanced for some functions such as q_x and m_x by the fourth difference error. A lack of compensation to this extent is therefore considered admissible. This formula becomes $\dfrac{[4][5][6]}{120}([3]-\gamma_2)f(x)$.

150. The most powerful summation formula which has been put to practical use is probably that of Spencer, which includes when expanded 21 terms of the original series. In this formula two summations in fives and one in sevens are used, so that $\dfrac{p^2+q^2+r^2-3}{24}=4$. Also $a=-\frac{1}{2}$, $b=0$, and $c=\frac{1}{2}$, so that the formula reduces to

$$\frac{[5]^2[7]}{175}(1+\tfrac{1}{2}\gamma_1-\tfrac{1}{2}\gamma_3)f(x)=\frac{[5]^2[7]}{350}(2+\gamma_1-\gamma_3)f(x)$$

$$=\frac{[5]^2[7]}{350}(1+[3]-\gamma_3)f(x).$$

151. A still more powerful formula would be given by putting $p=5$, $q=7$, and $r=11$, so that $\dfrac{p^2+q^2+r^2-3}{24}=8$, and

in the compensating adjustment we have $a=-1$, $b=0$, and $c=1$, so that the formula becomes,

$$\frac{[5][7][11]}{385}(1+\gamma_1-\gamma_3)f(x)=\frac{[5][7][11]}{385}([3]-\gamma_3)f(x).$$

152. The weight of a graduation formula is found by expanding the formula, adding together the squares of the coefficients and taking the reciprocal, that being the average proportion in which the mean square of the errors is reduced by the graduation. The smoothing coefficient is found by similarly expanding the third difference of the graduated terms, adding together the squares of the coefficients, dividing by twenty and extracting the square root. It measures the effect of the graduation on the mean absolute value of the third differences.

153. The following table shows a number of summation formulas and their weights and smoothing coefficients:

Author.	No. of Terms.	Formula.	Weight.	Smoothing Coefficient	Error.
Finlaison....	13	$\frac{[5]^3}{125}$	8.92	$\frac{1}{125}$	$3f''(x)+\frac{77}{20}f^{\text{IV}}(x)$
Woolhouse...	15	$\frac{[5]^3}{125}\{10-3[3]\}$	5.50	$\frac{1}{15}$	$-5.4f^{\text{IV}}(x)$
J. Spencer ..	15	$\frac{[5][4]^2}{320}\{4+3\gamma_1-3\gamma_2\}$	5.40	$\frac{1}{60}$	$-3.86f^{\text{IV}}(x)$
Higham.....	17	$\frac{[5]^3}{125}\{[3]-\gamma_2\}$	5.87	$\frac{1}{56}$	$-6.4f^{\text{IV}}(x)$
G. F. Hardy.	17	$\frac{[4][5][6]}{120}\{[3]-\gamma_2\}$	6.07	$\frac{1}{95}$	$\frac{1}{12}f''(x)-6.5f^{\text{IV}}(x)$
J. Spencer...	19	$\frac{[5]^2[7]}{175}\{3-\gamma_2\}$	6.73	$\frac{1}{85}$	$-10.6f^{\text{IV}}(x)$
Karup......	19	$\frac{[5]^3}{625}\{3[3]-2\gamma_3\}$	6.14	$\frac{1}{105}$	$-7.8f^{\text{IV}}(x)$
J. Spencer (*a*)	21	$\frac{[5]^2[7]}{350}\{1+[3]-\gamma_3\}$	6.98	$\frac{1}{160}$	$-12.6f^{\text{IV}}(x)$
J. Spencer (*b*)	21	$\frac{[4][5]^2[6]}{600}\{3-\gamma_2\}$	6.70	$\frac{1}{141}$	$\frac{1}{12}f''(x)-10.3f^{\text{IV}}(x)$
Henderson...	27	$\frac{[5][7][11]}{385}\{[3]-\gamma_3\}$	9.11	$\frac{1}{326}$	$-44.8f^{\text{IV}}(x)$

154. One difficulty in connection with the application of summation formulas to the graduation of tables is that in

order to determine any graduated value of the function it is necessary to know a number of ungraduated values above and below it and, unless the function is such that it disappears and the ungraduated values beyond a certain limit may be assumed to be zero, there will be a portion at either end of the table for which values will not be obtained, and some supplementary means must be adopted of completing the table if it is necessary to have it complete. In the case also of insurance companies' experiences, the values of the functions near the limits are usually derived from very limited data and are consequently very irregular. Both of these difficulties may be overcome by taking the three last graduated values of the function that are considered as determined with sufficient accuracy as a basis and determining a function of x of the third degree which will reproduce these three values and will also make the total expected deaths for ages beyond equal to the expected. The general expression for a function of the third degree in x such that $f(a)=u_a$, $f(b)=u_b$, and $f(c)=u_c$ is

$$f(x)=\frac{(x-b)(x-c)}{(a-b)(a-c)}u_a+\frac{(x-a)(x-c)}{(b-a)(b-c)}u_b+\frac{(x-a)(x-b)}{(c-a)(c-b)}u_c +k(x-a)(x-b)(x-c),$$

where k is an arbitrary constant.

155. Taking up next the application of a general law with arbitrary constants to the graduation of tables the law which is most frequently used is that of Makeham, according to which $\mu_x=A+Bc^x$, or $l_x=ks^xg^{c^x}$. The constant k is merely in the nature of a radix and does not affect the rates of mortality. This formula may be applied in two general ways: first, to construct a graduated mortality table from the original data of the exposed to risk (or population) and deaths without the explicit determination of the ungraduated rates of mortality; and second, to graduate a rough table without direct reference to the original data.

156. Probably the simplest method of determining the Makeham constants from the original data is to group the exposures (or population) and the deaths into quinquennial

age groups and then, by the process already described in connection with interpolation methods, determine the values of q_x, m_x or μ_x at quinquennial intervals. Where q_x or m_x is determined, we can proceed immediately to colog p_x from the known relations. Now we have

$$\begin{aligned}
\text{colog } p_x &= \log l_x - \log l_{x+1}, \\
&= (\log k + x \log s + c^x \log g) \\
&\qquad - \{\log k + (x+1) \log s + c^{x+1} \log g\}, \\
&= -\log s + (1-c)\, c^x \log g, \\
&= \alpha + \beta c^x.
\end{aligned}$$

It is therefore of the same form as μ_x. If, then, we neglect the values at young ages and at extreme old ages as derived from insufficient data and start with some age y in the neighborhood of age 30, we have

$$\begin{aligned}
&\text{colog } p_y + 3 \text{ colog } p_{y+5} + 5 \text{ colog } p_{y+10} + 5 \text{ colog } p_{y+15} \\
&\qquad + 3 \text{ colog } p_{y+20} + \text{colog } p_{y+25} = S_1 \\
&\qquad = 18\alpha + \beta c^y (1 + 3c^5 + 5c^{10} + 5c^{15} + 3c^{20} + c^{25}) \\
&\qquad = 18\alpha + \beta c^y (1 + c^5)(1 + c^5 + c^{10})^2.
\end{aligned}$$

Similarly

$$\begin{aligned}
&\text{colog } p_{y+15} + 3 \text{ colog } p_{y+20} + 5 \text{ colog } p_{y+25} + 5 \text{ colog } p_{y+30} \\
&\qquad + 3 \text{ colog } p_{y+35} + \text{colog } p_{y+40} = S_2 \\
&\qquad = 18\alpha + \beta c^{y+15} (1 + c^5)(1 + c^5 + c^{10})^2,
\end{aligned}$$

and

$$\begin{aligned}
&\text{colog } p_{y+30} + 3 \text{ colog } p_{y+35} + 5 \text{ colog } p_{y+40} + 5 \text{ colog } p_{y+45} \\
&\qquad + 3 \text{ colog } p_{y+50} + \text{colog } p_{y+55} = S_3 \\
&\qquad = 18\alpha + \beta c^{y+30} (1 + c^5)(1 + c^5 + c^{10})^2.
\end{aligned}$$

From these equations we see that

$$c^{15} = \frac{S_3 - S_2}{S_2 - S_1},$$

$$18\alpha = \frac{c^{15}S_1 - S_2}{c^{15} - 1} = \frac{S_1S_3 - S_2^2}{S_3 - 2S_2 + S_1},$$

$$\beta c^y(1+c^5)(1+c^5+c^{10})^2 = S_1 - 18\alpha = \frac{S_2 - S_1}{c^{15} - 1} = \frac{(S_2 - S_1)^2}{(S_3 - 2S_2 + S_1)}.$$

Thus, c, α, and β are determined. Similarly, the values of c, A, and B may be determined from the values of μ_x and from either α and β or A and B we may proceed to the values of s and g.

157. If a further refinement is required, we may assume that the values determined as above are only approximate and that

$$\text{colog } p_x = (\alpha + Mh) + (\beta + Mk)c^x\left(1 + \frac{l}{\beta}\right)^x,$$

where h, k, and l are small quantities and M is the modulus of common logarithms, or $\log_{10} e$. Then approximately

$$q_x = q'_x + hp'_x + kc^xp'_x + lxc^xp'_x,$$

where q'_x and p'_x are derived from the constants α, β, and c. We have, then, three unknowns, h, k, and l, to determine and three equations are required. These equations may be obtained by making the total number of expected deaths and the first and second moments of the expected deaths equal to those of the actual deaths.

The equations so obtained are:

$$h\Sigma E_xp'_x + k\Sigma c^xE_xp'_x + l\Sigma xc^xE_xp'_x = \Sigma(\theta_x - E_xq'_x),$$

$$h\Sigma xE_xp'_x + k\Sigma xc^xE_xp'_x + l\Sigma x^2c^xE_xp'_x = \Sigma x(\theta_x - E_xq'_x),$$

$$h\Sigma x^2E_xp'_x + k\Sigma x^2c^xE_xp'_x + l\Sigma x^3c^xE_xp'_x = \Sigma x^2(\theta_x - E_xq'_x).$$

These equations are seen to be equivalent to

$$h\Sigma E_xp'_x + (k+al)\Sigma c^xE_xp'_x + l\Sigma(x-a)c^xE_xp'_x = \Sigma(\theta_x - E_xq'_x),$$

$$h\Sigma(x-a)E_xp'_x + (k+al)\Sigma(x-a)c^xE_xp'_x + l\Sigma(x-a)^2c^xE_xp'_x$$
$$= \Sigma(x-a)(\theta_x - E_xq'_x),$$

$$h\Sigma(x-a)^2E_xp'_x + (k+al)\Sigma(x-a)^2c^xE_xp'_x + l\Sigma(x-a)^3c^xE_xp'_x$$
$$= \Sigma(x-a)^2(\theta_x - E_xq'_x),$$

where a is any suitable quantity used for the purpose of reducing the numbers involved. From these three equations the values of h, k, and l are determined, and the values of α, β, and c are corrected accordingly.

158. A shorter process is to assume that the value of c is accurate, and consequently $l=0$, and to determine h and k from the first two equations or to determine A and B directly from two equations depending on the relation

$$m_x=\mu_{x+\frac{1}{2}}=A+Bc^{x+\frac{1}{2}}.$$

The two equations are

$$A\Sigma L_x+B\Sigma c^{x+\frac{1}{2}}L_x=\Sigma\theta_x,$$

$$A\Sigma xL_x+B\Sigma xc^{x+\frac{1}{2}}L_x=\Sigma x\theta_x.$$

159. Where a mortality table has been constructed and it is desired to graduate it by Makeham's formula, the simplest method is to determine the constants from the values of $\log l_x$ at four equidistant ages by the method described in Chapter III. Unless, however, the ungraduated table is already comparatively smooth the constants so determined will depend to too great an extent on the particular ages selected. To minimize the irregularity we may take, instead of individual values of $\log l_x$, the sums of a number of consecutive values. Then we have:

$$S_1=\Sigma_x^{x+n-1}\log l_x=n\log k+\left\{nx+\frac{n(n-1)}{2}\right\}\log s+\frac{c^x(c^n-1)}{c-1}\log g,$$

$$S_2=\Sigma_{x+t}^{x+t+n-1}\log l_x=n\log k+\left\{n(x+t)+\frac{n(n-1)}{2}\right\}\log s+\frac{c^{x+t}(c^n-1)}{c-1}\log g,$$

$$S_3=\Sigma_{x+2t}^{x+2t+n-1}\log l_x=n\log k+\left\{n(x+2t)+\frac{n(n-1)}{2}\right\}\log s+\frac{c^{x+2t}(c^n-1)}{c-1}\log g,$$

$$S_4=\Sigma_{x+3t}^{x+3t+n-1}\log l_x=n\log k+\left\{n(x+3t)+\frac{n(n-1)}{2}\right\}\log s+\frac{c^{x+3t}(c^n-1)}{c-1}\log g,$$

$$S_2 - S_1 = nt \log s + \frac{c^x(c^t - 1)(c^n - 1)}{c - 1} \log g,$$

$$S_3 - S_2 = nt \log s + \frac{c^{x+t}(c^t - 1)(c^n - 1)}{c - 1} \log g,$$

$$S_4 - S_3 = nt \log s + \frac{c^{x+2t}(c^t - 1)(c^n - 1)}{c - 1} \log g,$$

$$S_3 - 2S_2 + S_1 = \frac{c^x(c^t - 1)^2(c^n - 1)}{c - 1} \log g,$$

$$S_4 - 2S_3 + S_2 = \frac{c^{x+t}(c^t - 1)^2(c^n - 1)}{c - 1} \log g,$$

$$\frac{S_4 - 2S_3 + S_2}{S_3 - 2S_2 + S_1} = c^t.$$

160. This method does not, however, entirely eliminate the objection that special importance is given to special points of division. To obviate this it has been suggested to use all the values of $\log l_x$ except those at extreme old ages and at young ages and to so determine the constants that the sum of the values and the first, second, and third moments will be reproduced. This can best be expressed in terms of summations as follows:

Suppose a is the youngest age to be included and let n be the total number of ages to be included. Then

$$\begin{aligned} \log l_{a+x} &= \log k + (a+x) \log s + c^{a+x} \log g, \\ &= (\log k + a \log s) + x \log s + c^x c^a \log g, \\ &= \log k' + x \log s + c^x \log g'. \end{aligned}$$

Where $\log k' = \log k + a \log s$ and $\log g' = c^a \log g$. Also

$$_1S_x = \Sigma_0^{x-1} \log {}_{a+x} = x \log k' + \frac{x^{(2)}}{2} \log s + \frac{c^x - 1}{c - 1} \log g',$$

$$_2S_x = \Sigma_0^{x-1} {}_1S_x = \frac{x^{(2)}}{2} \log k' + \frac{x^{(3)}}{6} \log s + \left\{ \frac{c^x - 1}{(c-1)^2} - \frac{x}{c-1} \right\} \log g',$$

$$_3S_x = \Sigma_0^{x-1} {}_2S_x = \frac{x^{(3)}}{6} \log k' + \frac{x^{(4)}}{24} \log s$$
$$+ \left\{ \frac{c^x - 1}{(c-1)^3} - \frac{x}{(c-1)^2} - \frac{x^{(2)}}{2(c-1)} \right\} \log g',$$

$$
{}_4S_x=\Sigma_0^{x-1}{}_3S_x=\frac{x^{(4)}}{24}\log k'+\frac{x^{(5)}}{120}\log s
$$

$$
+\left\{\frac{c^x-1}{(c-1)^4}-\frac{x}{(c-1)^3}-\frac{x^{(2)}}{2(c-1)^2}-\frac{x^{(3)}}{6(c-1)}\right\}\log g'.
$$

Where $x^{(r)}=x(x-1)(x-2)\ .\ .\ .\ (x-r+1)$.

From these four equations, after putting x equal to n in each, $\log k'$, $\log s$, and $\log g'$ may be eliminated, leaving an equation in c for solution. This equation will be of the $(n-1)$th degree and the numerical solution may be obtained to any required degree of accuracy. After the value of c is obtained those of $\log k'$, $\log s$, and $\log g'$ follow readily.

161. In the preceding discussion it has been assumed that the mortality table is an aggregate one, or in other words that it is analyzed only according to attained age, and where select or analyzed tables are required some modification of the method is necessary. This usually consists in making α and β, in the equation

$$
\text{colog } p_x=\alpha+\beta c^x,
$$

functions of the duration, so that we have as the general expression

$$
\text{colog } p_{[x]+t}=f_1(t)+f_2(t)c^{x+t}.
$$

The constants for each year of duration may be determined by any of the methods already described, the data for each year of duration being treated as representing a mortality table complete in itself but the same value of c being used for all.

162. The values of $f_1(t)$ and $f_2(t)$ so derived will, however, be somewhat irregular, so that they themselves require further graduation. It is usually assumed that they become constant after some definite duration, such as five years or ten years, the constant values being determined from the aggregate experience for all longer durations. In selecting formulas for graduating the values of these functions during the period of selection, the following conditions should be satisfied: (1) a smooth junction between the curves representing the select

and ultimate tables; (2) an agreement between the graduated and ungraduated values in the first year, as special importance is attached to the rate of mortality in that year; (3) an agreement between the aggregate graduated and ungraduated values of these functions during the period between the date of entry and the ultimate table. Considerable experimenting will usually be necessary to determine the function complying with these conditions.

163. In considering the method of graduation to be adopted in any particular case it is evident that a graduation by Makeham's formula possesses an advantage over the others on the scale of smoothness and since three arbitrary constants are available the sum of the deviations and of the accumulated deviations can be made to vanish. In view of these advantages and of its other advantages in connection with the calculation of joint contingencies, a graduation by that formula will be the best, provided the absolute deviations in groups do not materially exceed the expected and provided there is no characteristic feature of the experience which will not be reproduced by the formula.

164. Where a mathematical law cannot be applied it will usually be found that where the data is very scanty the graphic method will produce the best results as irregularities will occur of wide range, such as neither the interpolation nor the summation method is competent to remove. Where, however, the data is more extensive, so as to give a satisfactory degree of regularity under the operation of the interpolation or of the summation method, those methods will be the more satisfactory as the values derived do not depend on the judgment of the operator except as exercised in the selection of the particular graduation formula to be used, and they can be obtained to a greater degree of accuracy than is possible under the graphic method.

165. Diagram No. 2 illustrates the relation between an ungraduated series of rates of mortality and a graduated series. The irregular line represents the rate of mortality shown in the ungraduated experience, during the ten-year period from

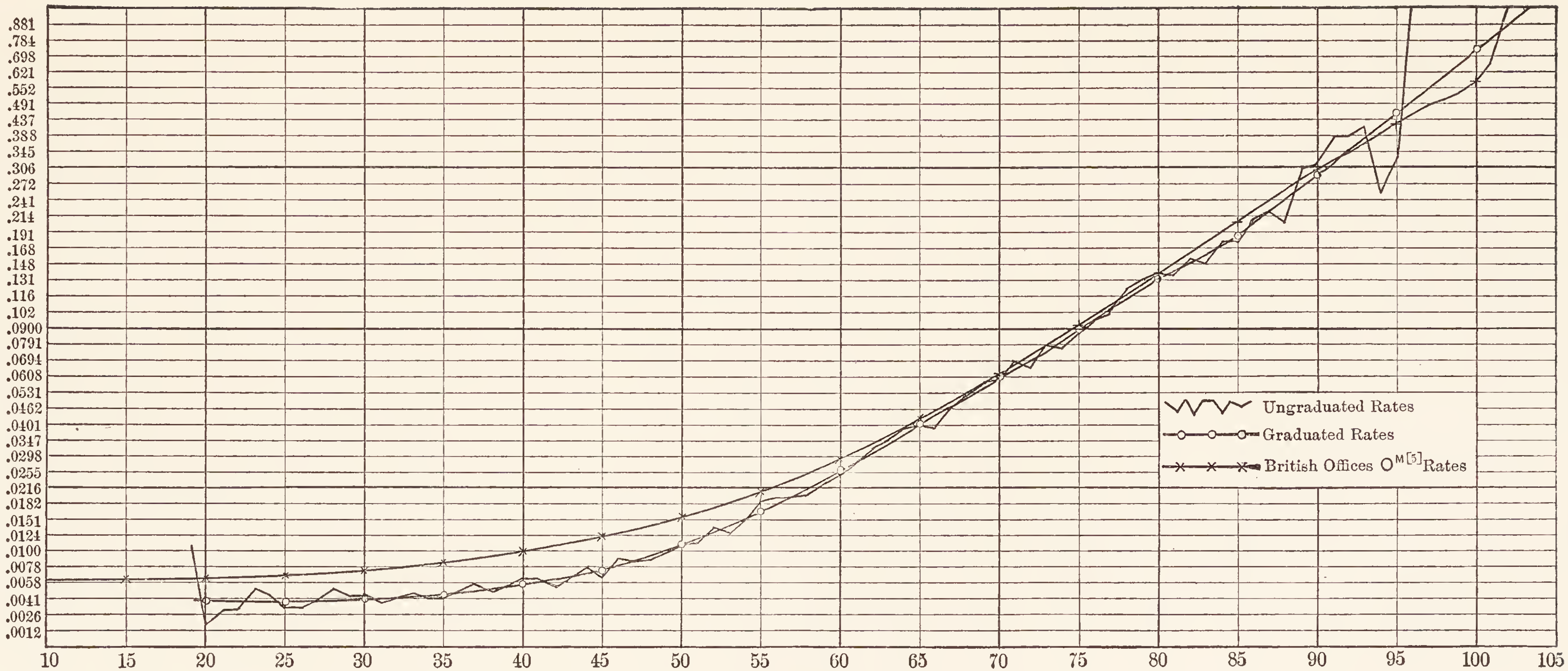

Diagram No. 2.—Comparison of Graduated and Ungraduated Rates of Mortality.

The Abscissas represent Ages. The Ordinates are proportional to values of log $(1+100\,q_x)$, where q_x is the probability of dying at age x, and the numbers on the left margin show values of q_x.

the policy anniversaries in 1899 to those in 1909, of a large American life insurance company on policies issued in the Northern States and in force more than five years at the time of observation. The ordinates are proportionate to log $(1+100q_x)$. The regular lines represent the same rates graduated by Spencer's formula with a preliminary adjustment at the extreme ages and the rates in the $O^{M(5)}$ table which was graduated by Makeham's formula. The rates by the M. A. table are not shown, partly because they would be practically indistinguishable in the diagram from those by the graduated experience of the company from age 35 to age 65 and from those by the $O^{M(5)}$ table from age 70 to age 100. The comparatively wide fluctuations in the rates by the ungraduated table at the extreme ages should be noted. The irregularity in the $O^{M(5)}$ table at the extreme old ages is due to the fact that the values of l_x and d_x were tabulated to integers only and the values of q_x recalculated from them instead of being calculated directly from the constants in the formula.

CHAPTER VIII

NORTHEASTERN STATES MORTALITY TABLE

166. Some of the methods described in the preceding chapters will be illustrated by the construction of a mortality table for the Northeastern States. The data used will be the death returns for the five calendar years 1908 to 1912 inclusive and the census returns as of June 1, 1900, and April 15, 1910, for the New England States and the three Middle Atlantic States, New York, New Jersey, and Pennsylvania. Table I shows the total deaths in these states for the five-year interval arranged by age groups and the total population at the two dates arranged in the same way. The average population by age groups for the five-year interval is also shown.

167. In this table the average population in each group is determined by means of Eqs. (5) and (9) of Chapter VI, only the population for which the ages are stated being taken into account. The cases where the age is returned as unknown are an extremely small percentage of the total and the effect of their omission is negligible. June 1, 1900, is $7\frac{7}{12}$ years before the beginning of the observations, which cover five years, therefore, $t_1 = -\frac{91}{60}$. April 15, 1910, is $2\frac{7}{24}$ years after the beginning therefore, $t_2 = \frac{11}{24}$. Also $P_3 = 21\,004\,724$ and $P_4 = 25\,836\,088$, so that we have $\log r = .0455239$ and the two factors entering into the determination of the average population are $-.03154363$ and 1.0304817 for June 1, 1900, and April 15, 1910, respectively. The total years of life are then obtained by multiplying the average population by five.

168. We then apply Eq. (3) of Chapter VII and obtain the values of $L_{9\frac{1}{2}}$, $L_{14\frac{1}{2}}$, $L_{18\frac{1}{2}}$, etc., and of $\theta_{14\frac{1}{2}}$, $\theta_{19\frac{1}{2}}$, $\theta_{24\frac{1}{2}}$, etc. The value of $\theta_{9\frac{1}{2}}$ is determined by leaving out of account the deaths

TABLE I

DEATHS AND POPULATION—NORTHEASTERN STATES. 1908-1912

Age Last Birthday.	Deaths, 1908–1912.	Population.			Total Years of Life.
		April 15, 1910	June 1, 1900.	Average 1908–1912.	
0	397 985	574 480	476 810	576 951	2 884 755
1	84 939	505 632	426 773	507 583	2 537 915
2	35 757	553 698	448 816	556 419	2 782 095
3	22 116	541 178	449 855	543 484	2 717 420
4	15 694	515 976	442 067	517 760	2 588 800
0–4	556 491	2 690 964	2 244 321	2 702 197	13 510 985
5–9	42 475	2 400 180	2 110 213	2 406 778	12 033 890
10–14	25 885	2 285 642	1 908 183	2 295 121	11 475 605
15–19	42 677	2 385 256	1 888 668	2 398 388	11 991 940
20–24	64 604	2 554 686	2 024 318	2 568 703	12 843 515
25–29	71 700	2 405 723	1 977 342	2 416 681	12 083 405
30–34	75 273	2 121 420	1 738 577	2 131 243	10 656 215
35–39	86 752	1 984 723	1 562 115	1 995 946	9 979 730
40–44	86 342	1 671 571	1 305 952	1 681 329	8 406 645
45–49	90 895	1 399 363	1 053 884	1 408 775	7 043 875
50–54	99 493	1 174 250	899 808	1 181 660	5 908 300
55–59	103 030	840 368	697 132	843 994	4 219 970
60–64	118 590	686 755	575 880	689 523	3 447 615
65–69	128 504	514 970	418 332	517 471	2 587 355
70–74	128 074	355 427	292 946	357 020	1 785 100
75–79	113 343	210 122	177 814	210 918	1 054 590
80–84	82 412	102 741	88 019	103 097	515 485
85–89	45 174	39 617	31 504	39 831	199 155
90–94	15 993	10 198	7 923	10 259	51 295
95–99	3 497	1 851	1 523	1 859	9 295
100 and over	678	261	270	260	1 300
All known	1 981 882	25 836 088	21 004 724	25 961 053	129 805 265
Unknown	1 038	32 485	41 971		
All	1 982 920	25 868 573	21 046 695		

at ages 0 to 4 and assuming that $\delta^2 w_7$ is equal to $\delta^2 w_{12}$ in the formula. The exposed to risk at each of these ages is then determined by the formula $E_x = L_x + \frac{1}{2}\theta_x$. The same formula also applies at ages 1 to 4 inclusive, but for age zero it is found that the average age at death of those dying within one year of birth is only three-tenths of a year, so that we use instead $E_0 = L_0 + \frac{7}{10}\theta_0$. From the values of log θ_x and

log E_x the values of log q_x are then determined. These figures are shown in Table II.

TABLE II

CALCULATION OF VALUES OF log q_x FOR INFANTILE AND QUINQUENNIAL AGES

x	$10L_x$	$10\theta_x$	$10E_x$	log q_x
0	28 847 550	3 979 850	31 633 445	$\bar{1}.09972$
1	25 379 150	849 390	25 803 845	$\bar{2}.51742$
2	27 820 950	357 570	27 999 735	.10621
3	27 174 200	221 160	27 284 780	$\bar{3}.90879$
4	25 888 000	156 940	25 966 470	.78134
$9\frac{1}{2}$	23 180 579	57 344	23 209 251	.39283
$14\frac{1}{2}$	23 234 918	62 207	23 266 021	.42712
$19\frac{1}{2}$	25 046 068	109 046	25 100 591	.63793
$24\frac{1}{2}$	25 302 916	138 167	25 372 000	.73605
$29\frac{1}{2}$	22 725 822	145 085	22 798 364	.80372
$34\frac{1}{2}$	20 660 018	162 848	20 741 442	.89494
$39\frac{1}{2}$	18 499 612	174 237	18 586 731	.97194
$44\frac{1}{2}$	15 378 331	175 751	15 466 206	$\bar{2}.05551$
$49\frac{1}{2}$	13 005 892	190 556	13 101 170	.16271
$54\frac{1}{2}$	10 068 338	201 374	10 169 025	.29672
$59\frac{1}{2}$	7 530 953	220 568	7 641 237	.46038
$64\frac{1}{2}$	6 039 903	249 733	6 164 770	.60756
$69\frac{1}{2}$	4 351 046	260 645	4 481 368	.76464
$74\frac{1}{2}$	2 796 270	246 450	2 919 495	.92642
$79\frac{1}{2}$	1 501 735	199 302	1 601 386	$\bar{1}.09501$
$84\frac{1}{2}$	650 085	127 131	713 651	.25077
$89\frac{1}{2}$	205 286	57 085	233 828	.38763
$94\frac{1}{2}$	37 512	15 140	45 082	.52612
$99\frac{1}{2}$	3 879	2 225	4 992	.64906

169. The value of log $q_{4\frac{1}{2}}$ to serve as an initial term in a systematic interpolation is calculated from those of log q_3, log q_4, and log $q_{9\frac{1}{2}}$, on the assumption that for the interval in question log q_x may be considered as a rational algebraic function of the second degree in x, the resulting value being $\bar{3}.72417$. An additional term is supplied at the end by assuming $q_{104\frac{1}{2}} = 1$ or log $q_{104\frac{1}{2}} = .00000$. Eq. (8) of Chapter VII is then applied to the interpolation. In this interpolation t is given the value 5 and h takes successively the values, $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$, $\frac{7}{2}$, and $\frac{9}{2}$. From these values of log q_x, the values of l_x and d_x are then derived. The following table for the first five ages shows the working process:

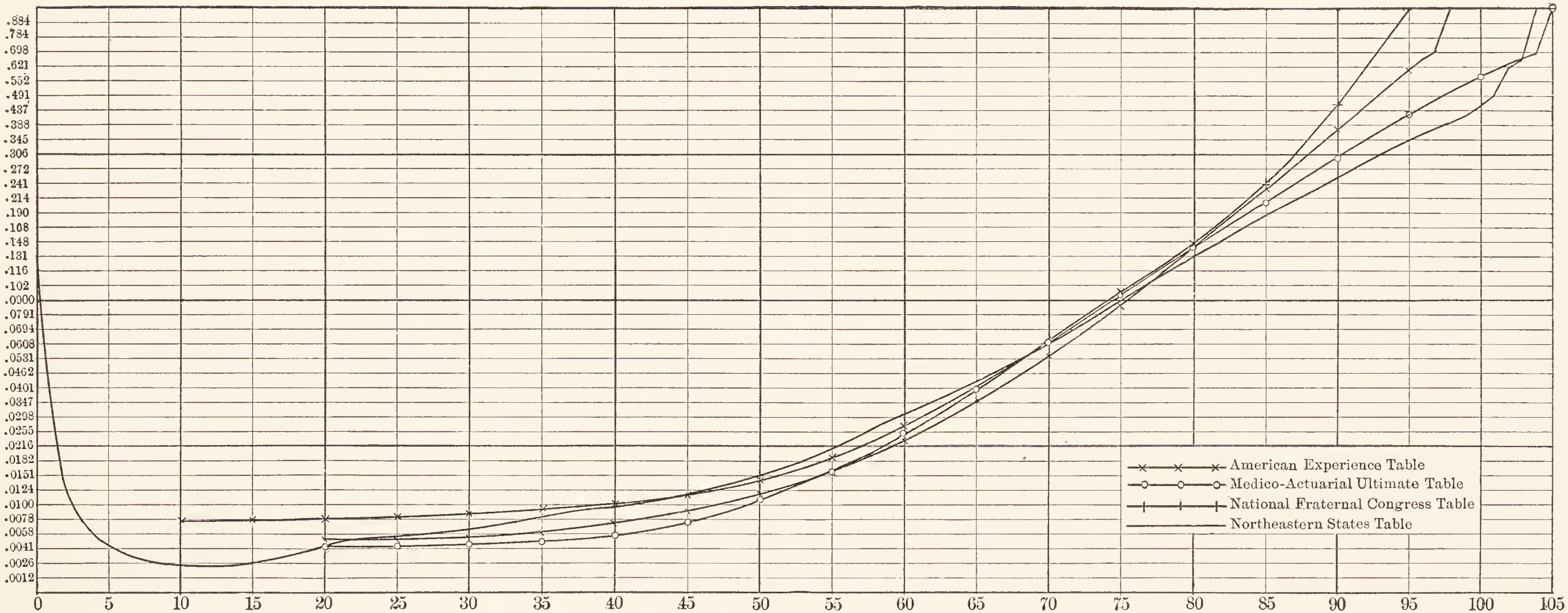

Diagram No. 3.—Rates of Mortality by Various American Tables of Mortality.

The Abscissas represent Ages. The Ordinates are proportional to the values of log $(1+100\,q_x)$, where q_x is the probability of dying at age x, and the numbers on the left margin show values of q_x.

(1) Age.	(2) l_x	(3) = log (2) log l_x	(4) log q_x	(5) = (3) + (4) log d_x	(6) = antilog (5) d_x
0	100 000	5.00000	$\bar{1}.09972$	4.09972	12 581
1	87 419	4.94161	$\bar{2}.51742$	3.45903	2 878
2	84 541	.92707	.10621	.03328	1 080
3	83 461	.92418	$\bar{3}.90879$	2.83027	677
4	82 784	.91795	.78134	.69929	500
5	82 284				

170. In accordance with the usual custom the values of q_x shown in the table have been adjusted to agree exactly with the values of l_x and d_x and do not agree with the values of log q_x used in constructing the table. In the accompanying Diagram, No. 3, the values of log $(1+100q_x)$ are plotted in comparison with the similar functions according to three tables representing mortality among American insured lives.

APPENDIX

Fundamental Columns and Other Data from Various Mortality Tables

x = age; l_x = number living at age x;
d_x = number dying at age x last birthday.

NORTHAMPTON TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
0	11 650	3 000	35	4 010	75	70	1 232	80
1	8 650	1 367	36	3 935	75	71	1 152	80
2	7 283	502	37	3 860	75	72	1 072	80
3	6 781	335	38	3 785	75	73	992	80
4	6 446	197	39	3 710	75	74	912	80
5	6 249	184	40	3 635	76	75	832	80
6	6 065	140	41	3 559	77	76	752	77
7	5 925	110	42	3 482	78	77	675	73
8	5 815	80	43	3 404	78	78	602	68
9	5 735	60	44	3 326	78	79	534	65
10	5 675	52	45	3 248	78	80	469	63
11	5 623	50	46	3 170	78	81	406	60
12	5 573	50	47	3 092	78	82	346	57
13	5 523	50	48	3 014	78	83	289	55
14	5 473	50	49	2 936	79	84	234	48
15	5 423	50	50	2 857	81	85	186	41
16	5 373	53	51	2 776	82	86	145	34
17	5 320	58	52	2 694	82	87	111	28
18	5 262	63	53	2 612	82	88	83	21
19	5 199	67	54	2 530	82	89	62	16
20	5 132	72	55	2 448	82	90	46	12
21	5 060	75	56	2 366	82	91	34	10
22	4 985	75	57	2 284	82	92	24	8
23	4 910	75	58	2 202	82	93	16	7
24	4 835	75	59	2 120	82	94	9	5
25	4 760	75	60	2 038	82	95	4	3
26	4 685	75	61	1 956	82	96	1	1
27	4 610	75	62	1 874	81			
28	4 535	75	63	1 793	81			
29	4 460	75	64	1 712	80			
30	4 385	75	65	1 632	80			
31	4 310	75	66	1 552	80			
32	4 235	75	67	1 472	80			
33	4 160	75	68	1 392	80			
34	4 085	75	69	1 312	80			

CARLISLE TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
0	10 000	1 539	35	5 362	55	70	2 401	124
1	8 461	682	36	5 307	56	71	2 277	134
2	7 779	505	37	5 251	57	72	2 143	146
3	7 274	276	38	5 194	58	73	1 997	156
4	6 998	201	39	5 136	61	74	1 841	166
5	6 797	121	40	5 075	66	75	1 675	160
6	6 676	82	41	5 009	69	76	1 515	156
7	6 594	58	42	4 940	71	77	1 359	146
8	6 536	43	43	4 869	71	78	1 213	132
9	6 493	33	44	4 798	71	79	1 081	128
10	6 460	29	45	4 727	70	80	953	116
11	6 431	31	46	4 657	69	81	837	112
12	6 400	32	47	4 588	67	82	725	102
13	6 368	33	48	4 521	63	83	623	94
14	6 335	35	49	4 458	61	84	529	84
15	6 300	39	50	4 397	59	85	445	78
16	6 261	42	51	4 338	62	86	367	71
17	6 219	43	52	4 276	65	87	296	64
18	6 176	43	53	4 211	68	88	232	51
19	6 133	43	54	4 143	70	89	181	39
20	6 090	43	55	4 073	73	90	142	37
21	6 047	42	56	4 000	76	91	105	30
22	6 005	42	57	3 924	82	92	75	21
23	5 963	42	58	3 842	93	93	54	14
24	5 921	42	59	3 749	106	94	40	10
25	5 879	43	60	3 643	122	95	30	7
26	5 836	43	61	3 521	126	96	23	5
27	5 793	45	62	3 395	127	97	18	4
28	5 748	50	63	3 268	125	98	14	3
29	5 698	56	64	3 143	125	99	11	2
30	5 642	57	65	3 018	124	100	9	2
31	5 585	57	66	2 894	123	101	7	2
32	5 528	56	67	2 771	123	102	5	2
33	5 472	55	68	2 648	123	103	3	2
34	5 417	55	69	2 525	124	104	1	1

ACTUARIES', OR COMBINED EXPERIENCE, TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
10	100 000	676	40	78 653	815	70	35 837	2 327
11	99 324	674	41	77 838	826	71	33 510	2 351
12	98 650	672	42	77 012	839	72	31 159	2 362
13	97 978	671	43	76 173	857	73	28 797	2 358
14	97 307	671	44	75 316	881	74	26 439	2 339
15	96 636	671	45	74 435	909	75	24 100	2 303
16	95 965	672	46	73 526	944	76	21 797	2 249
17	95 293	673	47	72 582	981	77	19 548	2 179
18	94 620	675	48	71 601	1 021	78	17 369	2 092
19	93 945	677	49	70 580	1 063	79	15 277	1 987
20	93 268	680	50	69 517	1 108	80	13 290	1 866
21	92 588	683	51	68 409	1 156	81	11 424	1 730
22	91 905	686	52	67 253	1 207	82	9 694	1 582
23	91 219	690	53	66 046	1 261	83	8 112	1 427
24	90 529	694	54	64 785	1 316	84	6 685	1 268
25	89 835	698	55	63 469	1 375	85	5 417	1 111
26	89 137	703	56	62 094	1 436	86	4 306	958
27	88 434	708	57	60 658	1 497	87	3 348	811
28	87 726	714	58	59 161	1 561	88	2 537	673
29	87 012	720	59	57 600	1 627	89	1 864	545
30	86 292	727	60	55 973	1 698	90	1 319	427
31	85 565	734	61	54 275	1 770	91	892	322
32	84 831	742	62	52 505	1 844	92	570	231
33	84 089	750	63	50 661	1 917	93	339	155
34	83 339	758	64	48 744	1 990	94	184	95
35	82 581	767	65	46 754	2 061	95	89	52
36	81 814	776	66	44 693	2 128	96	37	24
37	81 038	785	67	42 565	2 191	97	13	9
38	80 253	795	68	40 374	2 246	98	4	3
39	79 458	805	69	38 128	2 291	99	1	1

AMERICAN EXPERIENCE TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
10	100 000	749	40	78 106	765	70	38 569	2 391
11	99 251	746	41	77 341	774	71	36 178	2 448
12	98 505	743	42	76 567	785	72	33 730	2 487
13	97 762	740	43	75 782	797	73	31 243	2 505
14	97 022	737	44	74 985	812	74	28 738	2 501
15	96 285	735	45	74 173	828	75	26 237	2 476
16	95 550	732	46	73 345	848	76	23 761	2 431
17	94 818	729	47	72 497	870	77	21 330	2 369
18	94 089	727	48	71 627	896	78	18 961	2 291
19	93 362	725	49	70 731	927	79	16 670	2 196
20	92 637	723	50	69 804	962	80	14 474	2 091
21	91 914	722	51	68 842	1 001	81	12 383	1 964
22	91 192	721	52	67 841	1 044	82	10 419	1 816
23	90 471	720	53	66 797	1 091	83	8 603	1 648
24	89 751	719	54	65 706	1 143	84	6 955	1 470
25	89 032	718	55	64 563	1 199	85	5 485	1 292
26	88 314	718	56	63 364	1 260	86	4 193	1 114
27	87 596	718	57	62 104	1 325	87	3 079	933
28	86 878	718	58	60 779	1 394	88	2 146	744
29	86 160	719	59	59 385	1 468	89	1 402	555
30	85 441	720	60	57 917	1 546	90	847	385
31	84 721	721	61	56 371	1 628	91	462	246
32	84 000	723	62	54 743	1 713	92	216	137
33	83 277	726	63	53 030	1 800	93	79	58
34	82 551	729	64	51 230	1 889	94	21	18
35	81 822	732	65	49 341	1 980	95	3	3
36	81 090	737	66	47 361	2 070			
37	80 353	742	67	45 291	2 158			
38	79 611	749	68	43 133	2 243			
39	78 862	756	69	40 890	2 321			

INSTITUTE OF ACTUARIES HEALTHY MALE (H^M) TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
10	100 000	490	40	82 284	848	70	38 124	2 371
11	99 510	397	41	81 436	854	71	35 753	2 433
12	99 113	329	42	80 582	865	72	33 320	2 497
13	98 784	288	43	79 717	887	73	30 823	2 554
14	98 496	272	44	78 830	911	74	28 269	2 578
15	98 224	282	45	77 919	950	75	25 691	2 527
16	97 942	318	46	76 969	996	76	23 164	2 464
17	97 624	379	47	75 973	1 041	77	20 700	2 374
18	97 245	466	48	74 932	1 082	78	18 326	2 258
19	96 779	556	49	73 850	1 124	79	16 068	2 138
20	96 223	609	50	72 726	1 160	80	13 930	2 015
21	95 614	643	51	71 566	1 193	81	11 915	1 883
22	94 971	650	52	70 373	1 235	82	10 032	1 719
23	94 321	638	53	69 138	1 286	83	8 313	1 545
24	93 683	622	54	67 852	1 339	84	6 768	1 346
25	93 061	617	55	66 513	1 399	85	5 422	1 138
26	92 444	618	56	65 114	1 462	86	4 284	941
27	91 826	634	57	63 652	1 527	87	3 343	773
28	91 192	654	58	62 125	1 592	88	2 570	615
29	90 538	673	59	60 533	1 667	89	1 955	495
30	89 865	694	60	58 866	1 747	90	1 460	408
31	89 171	706	61	57 119	1 830	91	1 052	329
32	88 465	717	62	55 289	1 915	92	723	254
33	87 748	727	63	53 374	2 001	93	469	195
34	87 021	740	64	51 373	2 076	94	274	139
35	86 281	757	65	49 297	2 141	95	135	86
36	85 524	779	66	47 156	2 196	96	49	40
37	84 745	802	67	44 960	2 243	97	9	9
38	83 943	821	68	42 717	2 274			
39	83 122	838	69	40 443	2 319			

BRITISH OFFICES' $O^{M(5)}$ TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
10	107 324	658	45	82 010	984	80	15 531	2 151
11	106 666	658	46	81 026	1 018	81	13 380	2 007
12	106 008	656	47	80 008	1 056	82	11 373	1 847
13	105 352	655	48	78 952	1 096	83	9 526	1 674
14	104 697	654	49	77 856	1 139	84	7 852	1 493
15	104 043	654	50	76 717	1 185	85	6 359	1 308
16	103 389	654	51	75 532	1 234	86	5 051	1 122
17	102 735	655	52	74 298	1 286	87	3 929	943
18	102 080	655	53	73 012	1 343	88	2 986	773
19	101 425	655	54	71 669	1 402	89	2 213	617
20	100 770	657	55	70 267	1 464	90	1 596	480
21	100 113	660	56	68 803	1 529	91	1 116	360
22	99 453	661	57	67 274	1 598	92	756	263
23	98 792	664	58	65 676	1 669	93	493	183
24	98 128	667	59	64 007	1 742	94	310	124
25	97 461	672	60	62 265	1 819	95	186	79
26	96 789	676	61	60 446	1 897	96	107	49
27	96 113	681	62	58 549	1 975	97	58	28
28	95 432	688	63	56 574	2 055	98	30	15
29	94 744	694	64	54 519	2 133	99	15	8
30	94 050	703	65	52 386	2 211	100	7	4
31	93 347	711	66	50 175	2 285	101	3	2
32	92 636	720	67	47 890	2 355	102	1	1
33	91 916	732	68	45 535	2 421			
34	91 184	744	69	43 114	2 478			
35	90 440	757	70	40 636	2 527			
36	89 683	771	71	38 109	2 565			
37	88 912	788	72	35 544	2 591			
38	88 124	806	73	32 953	2 602			
39	87 318	825	74	30 351	2 596			
40	86 493	846	75	27 755	2 572			
41	85 647	869	76	25 183	2 529			
42	84 778	895	77	22 654	2 466			
43	83 883	922	78	20 188	2 381			
44	82 961	951	79	17 807	2 276			

NATIONAL FRATERNAL CONGRESS TABLE

x	l_x	d_x	x	l_x	d_x	x	l_x	d_x
20	100 000	500	50	81 702	935	80	20 270	2 799
21	99 500	501	51	80 767	981	81	17 471	2 659
22	98 999	502	52	79 786	1 029	82	14 812	2 485
23	98 497	503	53	78 757	1 083	83	12 327	2 280
24	97 994	505	54	77 674	1 140	84	10 047	2 050
25	97 489	507	55	76 534	1 202	85	7 997	1 800
26	96 982	510	56	75 332	1 270	86	6 197	1 539
27	96 472	513	57	74 062	1 342	87	4 658	1 277
28	95 959	517	58	72 720	1 418	88	3 381	1 023
29	95 442	522	59	71 302	1 501	89	2 358	788
30	94 920	527	60	69 801	1 588	90	1 570	579
31	94 393	533	61	68 213	1 681	91	991	404
32	93 860	540	62	66 532	1 778	92	587	264
33	93 320	548	63	64 754	1 880	93	323	161
34	92 772	557	64	62 874	1 985	94	162	89
35	92 215	567	65	60 889	2 094	95	73	44
36	91 648	578	66	58 795	2 206	96	29	19
37	91 070	591	67	56 589	2 318	97	10	7
38	90 479	606	68	54 271	2 430	98	3	3
39	89 873	622	69	51 841	2 539			
40	89 251	640	70	49 302	2 645			
41	88 611	660	71	46 657	2 744			
42	87 951	683	72	43 913	2 832			
43	87 268	708	73	41 081	2 909			
44	86 560	734	74	38 172	2 969			
45	85 826	761	75	35 203	3 009			
46	85 065	790	76	32 194	3 026			
47	84 275	822	77	29 168	3 016			
48	83 453	857	78	26 152	2 977			
49	82 596	894	79	23 175	2 905			

NORTHEASTERN STATES MORTALITY TABLE (1908-12)

Age. x	Number Living. l_x	Number Dying. d_x	Rate of Mortality per Thousand. $1000\, q_x$	Expectation of Life. $\mathring{e}_x$
0	100 000	12 581	125.81	50.41
1	87 419	2 878	32.92	56.59
2	84 541	1 080	12.77	57.50
3	83 461	677	8.11	57.24
4	82 784	500	6.04	56.70
5	82 284	386	4.69	56.04
6	81 898	311	3.80	55.31
7	81 587	261	3.20	54.51
8	81 326	228	2.80	53.69
9	81 098	207	2.55	52.84
10	80 891	195	2.41	51.97
11	80 696	190	2.35	51.10
12	80 506	190	2.36	50.21
13	80 316	196	2.44	49.33
14	80 120	207	2.58	48.45
15	79 913	222	2.78	47.58
16	79 691	244	3.06	46.71
17	79 447	270	3.40	45.85
18	79 177	299	3.78	45.00
19	78 878	329	4.17	44.17
20	78 549	354	4.51	43.36
21	78 195	374	4.78	42.55
22	77 821	390	5.01	41.75
23	77 431	402	5.19	40.96
24	77 029	413	5.36	40.17
25	76 616	424	5.53	39.38
26	76 192	435	5.71	38.60
27	75 757	445	5.87	37.82
28	75 312	456	6.05	37.05
29	74 856	468	6.25	36.26
30	74 388	482	6.48	35.49
31	73 906	499	6.75	34.72
32	73 407	518	7.06	33.95
33	72 889	537	7.37	33.19
34	72 352	557	7.70	32.43
35	71 795	575	8.01	31.68
36	71 220	591	8.30	30.93
37	70 629	607	8.59	30.18
38	70 022	623	8.90	29.44
39	69 399	639	9.21	28.70
40	68 760	656	9.54	27.96
41	68 104	674	9.90	27.23
42	67 430	693	10.28	26.49
43	66 737	713	10.68	25.76
44	66 024	734	11.12	25.04
45	65 290	758	11.61	24.31
46	64 532	785	12.16	23.59
47	63 747	813	12.75	22.88
48	62 934	845	13.43	22.17
49	62 089	879	14.16	21.46
50	61 210	915	14.95	20.76
51	60 295	955	15.84	20.07

NORTHEASTERN STATES MORTALITY TABLE (1908–12)—*Continued*

Age. x	Number Living. l_x	Number Dying. d_x	Rate of Mortality per Thousand. $1000q_x$	Expectation of Life. $\mathring{e}_x$
52	59 340	998	16.82	19.38
53	58 342	1 045	17.91	18.71
54	57 297	1 096	19.13	18.04
55	56 201	1 153	20.52	17.38
56	55 048	1 217	22.11	16.73
57	53 831	1 285	23.87	16.10
58	52 546	1 355	25.79	15.48
59	51 191	1 424	27.82	14.88
60	49 767	1 489	29.92	14.29
61	48 278	1 547	32.04	13.72
62	46 731	1 601	34.26	13.15
63	45 130	1 652	36.61	12.60
64	43 478	1 702	39.15	12.06
65	41 776	1 752	41.94	11.54
66	40 024	1 802	45.02	11.02
67	38 222	1 850	48.40	10.51
68	36 372	1 894	52.07	10.02
69	34 478	1 933	56.06	9.55
70	32 545	1 964	60.35	9.08
71	30 581	2 986	64.94	8.64
72	28 595	2 000	69.94	8.20
73	26 595	2 004	75.35	7.78
74	24 591	1 998	81.25	7.37
75	22 593	1 982	87.73	6.98
76	20 611	1 955	94.85	6.60
77	18 656	1 914	102.59	6.24
78	16 742	1 857	110.92	5.90
79	14 885	1 783	119.78	5.57
80	13 102	1 693	129.22	5.26
81	11 409	1 588	139.19	4.97
82	9 821	1 470	149.68	4.69
83	8 351	1 342	160.70	4.43
84	7 009	1 207	172.21	4.18
85	5 802	1 068	184.07	3.95
86	4 734	929	196.24	3.73
87	3 805	795	208.94	3.52
88	3 010	669	222.26	3.32
89	2 341	554	236.65	3.12
90	1 787	451	252.38	2.93
91	1 336	360	269.46	2.75
92	976	282	288.93	2.59
93	694	214	308.35	2.44
94	480	157	327.08	2.29
95	323	111	343.65	2.16
96	212	77	363.20	2.04
97	135	51	377.78	1.91
98	84	34	404.76	1.77
99	50	21	420.00	1.64
100	29	13	448.27	1.47
101	16	8	500.00	1.25
102	8	5	625.00	1.00
103	3	2	666.67	.83
104	1	1	1000.00	.50

RATES OF MORTALITY PER THOUSAND ACCORDING TO VARIOUS TABLES ($1000q_x$)

Age.	Northampton.	Carlisle.	English Life No. 3, Mixed.	Healthy Districts, Mixed.	English Life No. 6, Mixed.	North-eastern States.
0	257.51	153.90	149.49	102.95	156.53	125.81
5	29.44	17.80	13.43	10.27	7.43	4.69
10	9.16	4.49	5.73	4.36	2.50	2.41
15	9.22	6.19	5.36	4.87	3.11	2.78
20	14.03	7.06	8.42	7.30	4.36	4.51
25	15.76	7.31	9.38	8.08	5.36	5.53
30	17.10	10.10	10.32	8.57	6.48	6.48
35	18.70	10.26	11.42	9.03	8.36	8.01
40	20.91	13.00	12.87	9.69	10.84	9.54
45	24.02	14.81	14.84	10.82	13.17	11.61
50	28.35	13.42	17.53	12.62	17.12	14.95
55	33.50	17.92	22.76	15.35	22.77	20.52
60	40.24	33.49	30.66	22.99	32.32	29.92
65	49.02	41.09	43.46	35.35	45.44	41.94
70	64.94	51.64	63.80	53.24	67.20	60.35
75	96.15	95.52	93.94	80.54	95.59	87.73
80	134.33	121.72	135.51	120.43	146.00	129.22
85	220.43	175.28	189.29	174.95	215.41	184.07
90	260.87	260.56	254.84	245.03	290.56	252.38
95	750.00	233.33	331.17	325.11	396.91	343.65
100	....	222.22	412.56	413.04	600.00	448.27

Age.	Actuaries'.	American Experience.	Healthy Male.	British Offices, OM(5).	National Fraternal Congress.	Medico-Actuarial.
0						
5						
10	6.76	7.49	4.90	6.13		
15	6.94	7.63	2.87	6.29		
20	7.29	7.80	6.33	6.52	5.00	
25	7.77	8.06	6.63	6.89	5.20	4.7
30	8.42	8.43	7.72	7.47	5.55	4.9
35	9.29	8.95	8.77	8.37	6.15	5.1
40	10.36	9.79	10.31	9.78	7.17	5.7
45	12.21	11.16	12.19	12.00	8.87	7.5
50	15.94	13.78	15.95	15.45	11.44	10.6
55	21.66	18.57	21.03	20.83	15.71	15.8
60	30.34	26.69	29.68	29.21	22.75	24.0
65	44.08	40.13	43.43	42.21	34.39	39.0
70	64.93	61.99	62.19	62.19	53.65	61.7
75	95.56	94.37	98.36	92.67	85.48	91.9
80	140.41	144.47	144.65	138.50	138.09	137.2
85	205.10	235.55	209.89	205.69	225.08	203.7
90	323.73	454.55	279.45	300.75	368.79	297.8
95	584.27	1000.00	637.04	424.73	602.74	423.3
100	....	....	....	571.43	....	576.4

DEATH RATES PER THOUSAND ACCORDING TO VARIOUS TABLES

Table.	Infancy. Ages 0–2	Childhood. Ages 2–10.	Youth. Ages 10–25.	Maturity Ages 25–65.	Old Age. Ages over 65.
MIXED POPULATION TABLES:					
Northampton	249.31	32.49	11.60	24.23	91.87
Carlisle	130.33	24.21	6.26	15.30	89.89
English Life No. 3	118.19	16.07	6.79	16.47	89.45
Healthy Districts	73.48	10.93	5.89	12.66	81.37
English Life No. 6	116.71	8.57	3.65	14.96	92.35
Northeastern States	87.25	5.56	3.60	13.74	86.69
MALE POPULATION TABLES:					
English Life No. 3	128.24	16.10	6.62	17.03	92.35
Healthy Districts	80.28	10.85	5.43	12.76	83.31
English Life No. 6	128.07	8.60	3.74	16.24	96.79
FEMALE POPULATION TABLES:					
English Life No. 3	107.95	16.04	6.97	15.90	86.86
Healthy Districts	66.51	11.00	6.37	12.57	79.47
English Life No. 6	105.30	8.54	3.56	13.74	88.88
INSURANCE EXPERIENCE:					
Actuaries'	......	.....	7.14	14.86	91.12
Healthy Male H^M	......	.....	4.77	14.44	90.81
Healthy Male $H^{M(5)}$	......	.....	5.92	15.70	91.57
British Offices' O^M	......	.....	3.84	13.36	89.16
British Offices' $O^{M(5)}$	......	.....	6.42	14.14	89.25
American Experience	......	.....	7.99	13.65	90.12
National Fraternal Congress	......	.....	.....	10.92	84.50

INDEX